Speed of Vertex reinforced jump process on Galton-Watson trees

Xinxin Chen, Xiaolin Zeng

Speed of Vertex reinforced jump process on Galton-Watson trees

Xinxin Chen, Xiaolin Zeng

Abstract

We give an alternative proof of the fact that the vertex reinforced jump process on Galton-Watson tree has a phase transition between recurrence and transience as a function of c, the initial local time, see [3]. Further, applying techniques in [1], we show a phase transition between positive speed and null speed for the associated discrete time process in the transient regime.

1 Introduction and results

Let $\mathcal{G} = (V, E)$ be a locally finite graph endowed with its vertex set V and edge set E. Assign to each edge $e = \{u, v\} \in E$ a positive real number $W_e = W_{u,v}$ as its conductance, and assign to each vertex u a positive real number ϕ_u as its initial local time. Define a continuous-time V valued process $(Y_t; t \geq 0)$ on $\mathcal{G}$ in the following way: At time 0 it starts at some vertex $v_0 \in V$; If $Y_t = v \in V$, then conditionally on $\{Y_s; 0 \leq s \leq t\}$, the process jumps to a neighbor u of v at rate $W_{v,u} L_u(t)$ where

$$L_u(t) := \phi_u + \int_0^t \mathbf{1}_{\{Y_s = u\}} ds. \tag{1}$$

We call $(Y_t)_{t \geq 0}$ the vertex reinforced jump process (VRJP) on $(\mathcal{G}, W)$ starting from v_0.

It has been proved in [6] that when $\mathcal{G} = \mathbb{Z}$, (Y_t) is recurrent. When $\mathcal{G} = \mathbb{Z}^d$ with $d \geq 2$, the complete description of its behavior has not been revealed even though lots of effort has been made, see e.g. [2, 3, 5, 6, 7, 13].

Here we are interested in the case when $\mathcal{G}$ is a supercritical Galton-Watson tree, as we will see, acyclic property of trees largely reduces the difficulty to study this model. In [5] it is shown that the VRJP on 3-regular tree has positive speed and satisfies a central limit theorem. Later, Basdevant and Singh [3] gave a precise description of the phase transition of recurrence/transience for VRJP on supercritical Galton-Watson trees. In this paper, our main results, Theorem 2, describes the ballistic case of the VRJP when it is transient on supercritical Galton-Watson trees without leaves. Our proof is based on the random walk in random environment (RWRE) representation result of Sabot and Tarrès [13], and techniques on the studies of RWRE on trees, especially a result of Aidekon [1] (see also e.g.[**?**, **?**] for more on the studies of RWRE on trees).

Consider a rooted Galton-Watson tree T with offspring distribution $(q_k, k \geq 0)$ such that

$$b := \sum_{k \geq 0} k q_k > 1.$$

For some constant $c > 0$, we denote VRJP(c) the process (Y_t) on the Galton-Watson tree $T = (V, E)$ with $W_e \equiv 1$, $\forall e \in E$ and $\phi_x \equiv c$, $\forall x \in V$, starting from the root ρ. Hence the behaviors of this process depends on $\mathcal{G}$ and c. This definition is equivalent to VRJP with constant edge weight W and initial local time 1, up to a time change. We first recall the phase transition result obtained in [6]. Let A be an inverse Gaussian distribution of parameters $(1, c^2)$, i.e.

$$\mathbf{P}(A \in dx) = \mathbb{1}_{x\geq 0} \frac{c}{\sqrt{2\pi x^3}} \exp\Big\{ -\frac{c^2(x-1)^2}{2x} \Big\} dx, \tag{2}$$

The expectation w.r.t. $\mathbf{P}(dx)$ is denoted $\mathbf{E}$.

Theorem 1 (Basdevant & Singh)**.** *Let $\mu(c) = \inf_{a\in\mathbb{R}} \mathbf{E}[A^a] = \mathbf{E}[\sqrt{A}]$, then the VRJP($c$) on a supercritical GW tree with offspring mean b is recurrent a.s. if and only if $b\mu(c) \leq 1$.*

Remarks 1. *This phase transition was proved in [3] by considering the local times of VRJP. We will give another proof from the point of view of a random walk in random environment (RWRE), as a consequence of Theorem 3.*

When $b\mu(c) > 1$, a further question is to study the rate of escape of the process. Define the speed of the process (Y) by

$$v(Y) := \liminf_{t\to\infty} \frac{d(\rho, Y_t)}{t} = \lim_{t\to\infty} \frac{d(\rho, Y_t)}{t} \tag{3}$$

where d is the graph distance, and the last equality will be justified by Lemma 1. To study the speed, we use the RWRE point of view, relying on a result of Sabot & Tarrès [13], in particular, the following fact:

Let (Y_t) be a VRJP on a finite graph $\mathcal{G} = (V, E)$ with edge weight (W) and initial local time (ϕ). If (Z_t) is defined by

$$Z_t := Y_{D^{-1}(t)} \text{ where } D(t) := \sum_{x\in V} (L_x(t)^2 - \phi_x^2), \tag{4}$$

then (Z_t) is a mixture of Markov jump processes (c.f. also [14]). Moreover, the mixing measure is explicit.

Applying this result to our VRJP(c) on a tree, denote $(\eta_n)_{n\geq 0}$ the discrete time process associated to (Z_t), it turns out that (η_n) is a random walk in random environment. In [1], Aidekon gave a sharp and explicit criterion for the asymptotic speed to be positive, for random walks in random environment on Galton-Watson trees such that the environment is site-wise independent and identically distributed. This result cannot apply directly to the time changed VRJP(c), since the quenched transition probability depends also on the environment of the neighbors, see (7).

Aidekon's idea was to say that, most of the time the random walk will be wandering on long branches of the GW tree, it is then enough to look at the random walk on the half line. Thanks to the i.i.d. structure of the environment, he obtains sharp estimates for the one dimensional random walk, which allows him to come back to the tree without losing too much information. This also explains why the criterion depends on q_1, the probability that the GW tree generate one offspring.

In our case the environment is also i.i.d., the same idea will also work. Compare to [1], we mainly deal with the local dependences of the quenched probability transition. We believe that same type of criterion also holds for a larger type of random walk in random environment, with suitable conditions on the moments of the environment and locality of the transition probabilities.

Let us state our criterion, similar to (3), define

$$v(Z) = \liminf_{t\to\infty} \frac{d(\rho, Z_t)}{t},\ v(\eta) = \liminf_{n\to\infty} \frac{d(\rho, \eta_n)}{n}. \tag{5}$$

To study the speed, our techniques can only deal with trees without leaves, hence we assume that $q_0 = 0$. In addition, we assume that

$$M := \sum_{k\geq 0} k^2 q_k < \infty.$$

For any $r \in \mathbb{R}$, let

$$\xi_r = \xi_r(c) := \mathbf{E}[A^{-r}].$$

By (2), $\xi_r \in (0,\infty)$ for any r. In particular, $\mu(c) = \xi_{-1/2}(c)$. Our main theorem states that the speed depends on the value of q_1 and c.

Theorem 2. *Consider VRJP(c) on a supercritical GW tree such that $b\mu(c) > 1$, we have*

(1) $\lim_{t\to\infty} \frac{d(\rho,Z_t)}{t}$ and $\lim_{n\to\infty} \frac{d(\rho,\eta_n)}{n}$ exist almost surely,

(2) Assume $q_0 = 0$ and $M < \infty$. If $q_1\xi_{1/2} < 1$, then $v(\eta) > 0,\ v(Z) > 0$; if $q_1\xi_{1/2} > 1$, then $v(\eta) = v(Z) = 0$.

Corollary 1. *VRJP(c) $(Y_t)_{t\geq 0}$ on a supercritical GW tree such that $b\mu(c) > 1$, admits a speed $v(Y) \geq 0$ a.s. If in addition $q_0 = 0,\ M < \infty$ and $q_1\xi_{1/2} < 1$, then $v(Y) > 0$.*

Remarks 2. *Our method cannot tackle the critical case $q_1\xi_{1/2} = 1$. Moreover, whether $q_1\xi_{1/2} > 1$ implies $v(Y) = 0$ remains unknown.*

The rest of this paper is organized as follows. In Section 2, we use a result of Sabot & Tarres [13] to recover the RWRE structure of VRJP. Section 3 is devoted to an alternative proof of Theorem 1, as an application of the RWRE point of view. Section 4 establishes the existence of the speed for the RWRE and Theorem 2. The proofs of some technical lemmas are left in Appendix.

2 RWRE on Galton-Watson tree

2.1 Mixture of Markov jump process by changing times

In this subsection, we consider a VRJP $(Y_t)_{t\geq 0}$ on a tree $T = (V,E)$ rooted at ρ, with edge weights (W) and initial local time (ϕ). If $x \neq \rho$, let $\overleftarrow{x}$ be the parent of x on the tree, the associated edge is denoted by $e_x = (x, \overleftarrow{x})$ with weight W_{e_x}.

Recall that the time changed version of VRJP (Z_t) defined in (4) is mixture of Markov jump processes with correlated mixing measure. The advantage of considering VRJP on trees is that, the random environment becomes independent.

Theorem 3. *Let $T = (V,E)$ be a tree rooted at ρ, endowed with edge weights $(W_e)_{e\in E}$ and initial local times $(\phi_x)_{x\in V}$. Let $(A_x, x \in V\setminus\{\rho\})$ be independent random variables defined by*

$$\mathbf{P}(A_x \in da) = \mathbb{1}_{\mathbb{R}^+}(a)\phi_x\sqrt{\frac{W_{e_x}\phi_x\phi_{\overleftarrow{x}}}{2\pi a^3}}\exp(-W_{e_x}\phi_x\phi_{\overleftarrow{x}}\frac{(a-1)^2}{2a})da.$$

If X_t is a mixture of Markov jump processes starting from ρ, such that, conditionally on $(A_x, x \in V\setminus\{\rho\})$, X_t jumps from x to $\overleftarrow{x}$ at rate $\frac{1}{2}W_{e_x}\frac{\phi_{\overleftarrow{x}}}{\phi_x A_x}$ and from $\overleftarrow{x}$ to x at rate $\frac{1}{2}W_{e_x}\frac{\phi_x A_x}{\phi_{\overleftarrow{x}}}$. Then X_t and Z_t (defined in (4)) has the same distribution.

Proof. On trees, VRJP observed at times when it stays on any finite sub-tree $T_f = (V_f, E_f)$ (also rooted at ρ) of T, behaves the same as VRJP restricted to T_f; moreover, the restriction is independent of the VRJP outside T_f. Therefore, it is enough to prove the theorem on finite tree T_f. By Theorem 2 of [13] (with a slight modification of the initial local time, or a more detailed version in [**?**], appendix B), if we denote

$$l_x(t) = \int_0^t \mathbb{1}_{Z_s = x} ds,$$

then

$$U_x = \frac{1}{2} \lim_{t\to\infty} \left(\log \frac{l_x(t) + \phi_x^2}{l_\rho(t) + \phi_\rho^2} - \log \frac{\phi_x^2}{\phi_\rho^2} \right)$$

exists a.s. and $\{U_x, x \in V_f,\ U_\rho = 0\}$ has distribution (where $du = \prod_{x\neq\rho} du_x$)

$$d\mathcal{Q}_{\rho,T_f}^{W,\phi}(u) = \frac{\prod_{x\neq\rho}\phi_x}{\sqrt{2\pi}^{|V_f|-1}} e^{-\sum_{x\in V_f} u_x - \sum_{\{x,y\}\in E_f} \frac{1}{2} W_{x,y}(e^{u_x-u_y}\phi_y^2 + e^{u_y-u_x}\phi_x^2 - 2\phi_x\phi_y)} \sqrt{\prod_{\{x,y\}\in E_f} W_{x,y} e^{u_x+u_y}} du.$$

Now, conditionally on (U_x), Z_t is a Markov process which jumps at rate (from x to z) $\frac{1}{2}W_{x,z}e^{U_z-U_x}$. For $e_x = (x, \overleftarrow{x}) \in T_f$, if we writes $y_{e_x} = (u_{\overleftarrow{x}} - \log\phi_{\overleftarrow{x}}) - (u_x - \log\phi_x)$, then (note that $u \mapsto y$ is a diffeomorphism and $dy = du$) the density of (u) also writes

$$d\mathcal{Q}_{\rho,T_f}^{W,\phi}(u) = \prod_{e_x = \{x,\overleftarrow{x}\}\in E_f} \sqrt{\frac{W_{e_x}\phi_x\phi_{\overleftarrow{x}}}{2\pi}} \exp\left(\frac{1}{2}(y_{e_x} - W_{e_x}\phi_x\phi_{\overleftarrow{x}}(e^{y_{e_x}} + e^{-y_{e_x}} - 2)) \right) dy.$$

Plugging $a_x = e^{-y_{e_x}}$ entails that a_x is Inverse Gaussian distributed with parameter $(1, W_{e_x}\phi_x\phi_{\overleftarrow{x}})$ and

$$d\mathcal{Q}_{\rho,T_f}^{W,\phi}(a) = \prod_{x\in V_f\setminus\{\rho\}} \mathbb{1}_{a_x>0} \sqrt{\frac{W_{e_x}\phi_x\phi_{\overleftarrow{x}}}{2\pi a_x^3}} \exp(-W_{e_x}\phi_x\phi_{\overleftarrow{x}} \frac{(a_x-1)^2}{2a_x}) da_x$$

Finally note that

$$\frac{1}{2} W_{x,z} e^{u_z - u_x} = \begin{cases} \frac{1}{2} W_{x,z} \frac{\phi_z}{\phi_x a_x} & \text{if } z = \overleftarrow{x} \\ \frac{1}{2} W_{x,z} \frac{\phi_z a_z}{\phi_x} & \text{if } \overleftarrow{z} = x. \end{cases}$$

□

For VRJP(c) on a GW tree, the theorem immediately implies:

Corollary 2. *On a sampled GW tree $T = (V, E)$, the time changed VRJP(c) (Z_t) is a random walk in i.i.d. environment $(A_x, x \in V \setminus \{\rho\})$, where (A_x) are i.i.d. inverse Gaussian distributed with parameters $(1, c^2)$, and conditionally on the environment, the process jumps at rate*

$$\begin{cases} \frac{1}{2A_x} & \text{from } x \text{ to } \overleftarrow{x} \\ \frac{1}{2}A_x & \text{from } \overleftarrow{x} \text{ to } x. \end{cases} \tag{6}$$

2.2 RWRE on Galton Watson tree and notations

In the sequel, let $T = (V, E)$ be a Galton-Watson tree with offspring distribution $\{q_k; k \geq 0\}$. Recall that $(\eta_n)_{n\geq 0}$ denotes the discrete time process associated to (Z_t) (or (Y_t)), which is a random walk in random environment.

Note that there are two level of randomnesses in the environment. First, we sample a GW tree, T, whose law is denoted by $GW(dT)$. Then, given the tree T (rooted at ρ), we define $\omega = \{A_x, x \in V \setminus \{\rho\}\}$ as in Corollary 2, whose law is $\prod_{x\in T\setminus\{\rho\}} \mathbf{P}(dA_x)$, which we denote abusively $\mathbf{P}(d\omega)$. Finally, given (w, T), the Markov jump process $(Z_t; t \geq 0)$ is defined by its jump rate in (6).

For convenience, we artificially add a vertex $\overleftarrow{\rho}$ to T, designing the parent of the root. Let A_ρ be another copy of A, independent of all others. Now, (abusively) let $\omega = (A_x, x \in V)$ be the enlarged environment. Given (ω, T), define the new Markov chain η, which is a random walk on $V \cup \{\overleftarrow{\rho}\}$, with transition probabilities

$$\begin{cases} p(x, \overleftarrow{x}) = \frac{1}{1+A_x \sum_{y:\overleftarrow{y}=x} A_y} \\ p(x, z) = \frac{A_x A_z}{1+A_x \sum_{y:\overleftarrow{y}=x} A_y} \quad \text{where } \overleftarrow{z} = x \in V \\ p(\overleftarrow{\rho}, \rho) = 1 \end{cases} \tag{7}$$

This modification will not change the recurrence/transience behavior of the RWRE η nor its speed in the transient regime. We will always work with this modification in the sequel.

Let us now introduce the notation of quenched and annealed probabilities. Given the environment (ω, T), let $P_x^{\omega,T}$ denote the quenched probability of the random walk η with $\eta_0 = x \in V$ a.s. Denote by $\mathbb{P}_x^T$, $\mathbf{Q}$, $\mathbb{P}_\rho$ the mesures:

$$\begin{aligned} \mathbb{P}_x^T(\cdot) &:= \int P_x^{\omega,T}(\cdot)\mathbf{P}(d\omega), \\ \mathbf{Q}(\cdot) &:= \int \mathbb{1}_{\{\cdot\}}\mathbf{P}(d\omega)GW(dT) \\ \mathbb{P}_\rho(\cdot) &:= \int \mathbb{P}_\rho^T(\cdot)GW(dT), \end{aligned}$$

and the associated expectations are denoted $E_x^{\omega,T}$, $\mathbb{E}_x^T$, $\mathbb{E}_{\mathbf{Q}}$ and $\mathbb{E}$. Note the slight difference for the expectation corresponds to $\mathbf{Q}$: $\mathbb{E}_{\mathbf{Q}}$. For brevity, we omit the starting point if the random walk starts from the root; that is, we write $P^{\omega,T}$, $\mathbb{P}^T$ and $\mathbb{P}$ for $P_\rho^{\omega,T}$, $\mathbb{P}_\rho^T$ and $\mathbb{P}_\rho$, Notice that $\mathbb{P}$ is the annealed law of η.

For any vertex x, let $|x| = d(\rho, x)$ be the generation of x and denote by $[\![\rho, x]\!]$ the unique shortest path from x to the root ρ, and x_i (for $0 \leq i \leq |x|$) the vertices on $[\![\rho, x]\!]$ such that $|x_i| = i$. In particular, $x_0 = \rho$ and $x_{|x|} = x$. In words, x_i (for $i < |x|$) is the ancestor of x at generation i. Also denote $]\!]\rho, x]\!] := [\![\rho, x]\!]\setminus\{\rho\}$ and $]\!]\rho, x[\![:= [\![\rho, x]\!]\setminus\{\rho, x\}$.

3 Phase transition: an alternative proof of Theorem 1

The ideas follow from Lyons and Pemantle [11], by means of random electrical network.

Proof of Theorem 1. Recall that the environment ω is given by i.i.d. random variables A_x, $x \in T$, with inverse Gaussian distribution $IG(1, c^2)$. The RWRE is equivalent to an electrical network with random conductances:

$$C_{e_x} := C(x, \overleftarrow{x}) = (\prod_{u\in]\!]\rho, x[\![} A_u)^2 A_x, \forall x \in V \setminus \{\rho\}.$$

We omit the proof of the transient case which is quite similar to that in Lyons and Pemantle [11], however, we will detail the recurrence case. That is, we will show that if $b\mu(c) \leq 1$, then the RWRE is recurrent a.s.

First consider the case $b\mu(c) < 1$, note that

$$\begin{aligned}\mathbb{E}_{\mathbf{Q}}\Big[\sum_{n\geq 1}\sum_{|x|=n} C_{e_x}^{1/4}\Big] &= \sum_{n\geq 1}\int\Big(\int\sum_{|x|=n} C_{e_x}^{1/4}\mathbf{P}(d\omega)\Big)GW(dT)\\ &= \sum_{n\geq 1}\int\sum_{|x|=n}\mathbf{E}[A^{1/2}]^{n-1}\mathbf{E}[A^{1/4}]GW(dT)\\ &= \sum_{n\geq 1} b^n\mathbf{E}[A^{1/2}]^{n-1}\mathbf{E}[A^{1/4}].\end{aligned}$$

Because $\mu(c) = \mathbf{E}[A^{1/2}] < 1/b$, we have, for some constants $c_1, c_2 \in \mathbb{R}^+$

$$\mathbb{E}_{\mathbf{Q}}\Big[\sum_{n\geq 1}\sum_{|x|=n} C_{e_x}^{1/4}\Big] \leq c_1\sum_{n\geq 0}(b\mu(c))^n \leq c_2 < \infty,$$

which implies that

$$\sum_{n\geq 1}\sum_{|x|=n} C_{e_x}^{1/4} < \infty,\ \mathbf{Q}\text{-a.s.}$$

As a result, there exists a stationary probability a.s., moreover η is positive recurrent.

Turning to the case $b\mu(c) = 1$, let $\Pi_n := \{e_x : |x| = n\}$ be a sequence of cutsets. Observe that

$$W_n := \sum_{|x|=n}\prod_{u\in\,]\!]\rho,x]\!]} A_u^{1/2} = \sum_{|x|=n} C_{e_x}^{1/4}A_x^{1/4}.$$

is a martingale with respect to its natural filtration. By Biggin's theorem ([4, 10]), it converges a.s. to zero. We are going to show that $\mathbf{Q}$-a.s.,

$$\liminf_{n\to\infty}\sum_{|x|=n} C_{e_x}^{1/4} = 0, \tag{8}$$

in particular, this will imply that $\mathbf{Q}$-a.s. $\inf_{\Pi:\text{ cutset}}\sum_{e_x\in\Pi} C_{e_x} = 0$. By the trivial half of the max-flow min-cut theorem, the corresponding network admits no flow a.s. Hence, the random walk is a.s. recurrent. Observes that

$$\begin{aligned}\sum_{|x|=n} C_{e_x}^{1/4} &= \sum_{|x|=n}\prod_{u\in\,]\!]\rho,x[\![} A_u^{1/2}A_x^{1/4}1_{\{A_x\geq 1\}} + \sum_{|x|=n}\prod_{u\in\,]\!]\rho,x[\![} A_u^{1/2}A_x^{1/4}1_{\{A_x<1\}}\\ &= \sum_{|x|=n}\prod_{u\in\,]\!]\rho,x]\!]} A_u^{1/2}A_x^{-1/4}1_{\{A_x\geq 1\}} + \sum_{|y|=n-1}\prod_{u\in\,]\!]\rho,y]\!]} A_u^{1/2}\sum_{x:\overleftarrow{x}=y} A_x^{1/4}1_{\{A_x<1\}}\\ &\leq W_n + \sum_{|y|=n-1}\prod_{u\in\,]\!]\rho,y]\!]} A_u^{1/2}\nu_y,\end{aligned}$$

where ν_y denotes the number of children of y. Letting n go to infinity yields that

$$0 \leq \liminf_{n\to\infty}\sum_{|x|=n} C_{e_x}^{1/4} \leq \liminf_{n\to\infty}\sum_{|y|=n-1}\prod_{u\in\,]\!]\rho,y]\!]} A_u^{1/2}\nu_y.$$

For any $K \geq 1$, separating the sum over vertices y according to $\{\nu_y < K\}$ or $\{\nu_y \geq K\}$, the last term is bounded by

$$\begin{aligned}&\lim_{n\to\infty} KW_{n-1} + \liminf_{n\to\infty}\sum_{|y|=n-1}\prod_{u\in\,]\!]\rho,y]\!]} A_u^{1/2}\nu_y 1_{\{\nu_y\geq K\}}\\ &= \liminf_{n\to\infty}\sum_{|y|=n-1}\prod_{u\in\,]\!]\rho,y]\!]} A_u^{1/2}\nu_y 1_{\{\nu_y\geq K\}}.\end{aligned}$$

By Fatou's lemma,

$$\begin{aligned}
&\mathbb{E}_{\mathbf{Q}}\Big(\liminf_{n\to\infty}\sum_{|y|=n-1}\prod_{u\in\,]\!]\rho,\,y]\!]}A_u^{1/2}\nu_y 1_{\{\nu_y\geq K\}}\Big)\\
&\leq \liminf_{n\to\infty}\mathbb{E}_{\mathbf{Q}}\Big(\sum_{|y|=n-1}\prod_{u\in\,]\!]\rho,\,y]\!]}A_u^{1/2}\nu_y 1_{\{\nu_y\geq K\}}\Big)=\mathbb{E}_{\mathbf{Q}}[\nu_\rho,\nu_\rho\geq K],
\end{aligned}$$

since for all $|y|=n-1$, ν_y is independent of $\prod_{u\in\,]\!]\rho,\,y]\!]}A_u^{1/2}$ and $\mathbb{E}_{\mathbf{Q}}\Big(\sum_{|y|=n-1}\prod_{u\in\,]\!]\rho,\,y]\!]}A_u^{1/2}\Big)=1$. Consequently, for any $K\geq 1$,

$$\mathbb{E}_{\mathbf{Q}}\Big[\liminf_{n\to\infty}\sum_{|x|=n}C_{e_x}^{1/4}\Big]\leq \mathbb{E}_{\mathbf{Q}}[\nu_\rho,\nu_\rho\geq K].$$

As $b=\mathbb{E}_{\mathbf{Q}}[\nu_y]<\infty$, letting $K\to\infty$ gives

$$\mathbb{E}_{\mathbf{Q}}\Big[\liminf_{n\to\infty}\sum_{|x|=n}C_{e_x}^{1/4}\Big]=0.$$

This implies (8). □

4 Speed when transient

Turning to the positivity of $v(Z)$ and $v(\eta)$, note that the processes (Z_t) and (η_n) are mixture of Markov processes but (Y_t) is not, in fact, (Y_t) escapes faster than (Z_t), in particular, when $v(Z)>0$, we have $v(Y)>0$. But we are not sure whether $v(Z)=0$ implies $v(Y)=0$.

4.1 Regeneration structure

In this section, we show that, when the process (η_n) (or (Z_t)) is transient, its path can be cut into independent pieces, using the notion of regeneration time. As a consequence, the speed $v(\eta)$, $v(Z)$ exists a.s. as a limit (not just a $\liminf$).

On a tree, when a random walk traverses an edge for the first and last time simultaneously, we say it regenerates since it will now remain in a previously unexplored sub-tree. For any vertex x, let $\mathcal{D}(x)=\inf\{k\geq 1,\ \eta_{k-1}=x,\eta_k=\overleftarrow{x}\}$, write $\tau_n=\inf\{k\geq 0,\ |\eta_k|=n\}$ and define the regeneration time recursively by

$$\begin{cases}\Gamma_0=0\\ \Gamma_n=\Gamma_n(\eta)=\inf\{k>\Gamma_{n-1};\ d(\eta_k)\geq 3,\mathcal{D}(\eta_k)=\infty,\tau_{|\eta_k|}=k\}.\end{cases}$$

where $d(x)$ is the degree of the vertex x.

Lemma 1. *Let $S(\cdot)=\mathbb{P}(\cdot|d(\rho)\geq 3,\ \mathcal{D}(\rho)=\infty)$, if η is transient, then*

i) For any $n\geq 1$, $\Gamma_n<\infty$ $\mathbb{P}$-a.s.

ii) Under $\mathbb{P}$, $(\Gamma_{n+1}-\Gamma_n,|\eta_{\Gamma_{n+1}}|-|\eta_{\Gamma_n}|,A_{\Gamma_{n+1}})_{n\geq 1}$ are independent and distributed as $(\Gamma_1,|\eta_{\Gamma_1}|,A_{\Gamma_1})$ under S.

iii) $E_S(|\eta_{\Gamma_1}|)<\infty$.

We feel free to omit the proof because it is analogue to 'Fact' in [1] p.10. In addition, Lemma 1 also holds without assuming $d(\eta_k) \geq 3$ in the definition of Γ_n, but we will need this assumption later in the proof of Lemma 7.

By strong law of large numbers, one immediately sees that there exist two constants $c_4 \geq c_3 \geq 1$ such that $\mathbb{P}$-a.s.,

$$\lim_{n\to\infty} \frac{|\eta_{\Gamma_n}|}{n} = c_3 \in [1, \infty), \quad \lim_{n\to\infty} \frac{\Gamma_n}{n} = c_4 \in [c_3, \infty].$$

In addition, for any $n \geq 1$, there exists a unique $u(n) \in \mathbb{N}$ such that

$$\Gamma_{u(n)} \leq n < \Gamma_{u(n)+1}$$

and $|\eta_{\Gamma_{u(n)}}| \leq |\eta_n| < |\eta_{\Gamma_{u(n)+1}}|$. Letting n go to infinity, (in particular $u(n) \to \infty$) in

$$\frac{|\eta_{\Gamma_{u(n)}}|}{\Gamma_{u(n)+1}} \leq \frac{|\eta_n|}{n} < \frac{|\eta_{\Gamma_{u(n)+1}}|}{\Gamma_{u(n)}} = \frac{|\eta_{\Gamma_{u(n)+1}}|}{u(n)} \frac{u(n)}{\Gamma_{u(n)}}.$$

We have $\mathbb{P}$-a.s.

$$\frac{|\eta_n|}{n} \to v(\eta) := \frac{c_3}{c_4} \in [0, 1].$$

For Z_t, the same arguments can be applied. As a consequence of the i.i.d. decomposition, $v(Z) = \lim_{t\to\infty} \frac{|Z_t|}{t}$ exists a.s. The existence of $v(Y) = \lim_{t\to\infty} \frac{|Y_t|}{t}$ can be justified by performing the time change $D(t)$ between consecutive regenerative epochs.

4.2 The auxiliary one dimensional process

The RWRE can also be defined on the deterministic graph $\mathbb{H} = \{-1, 0, 1, \ldots\}$, on which many quantities are viable by explicit computations. The strategy is to compare the random walk on a tree to the random walk on the half line, in the forth coming sections we will explain how these comparisons will be done. In this section we list some properties of the one dimensional random walk, their proofs can be found in Appendix A.

Let $\tilde{\eta}_n$ be the random walk on the half line $\mathbb{H} = \{-1, 0, 1, \ldots\}$ in the random environment $\omega = (A_k, k \geq 0)$ which are i.i.d. copies of A under $\mathbf{P}$, with transition probability according to (7); that is,

$$\begin{cases} p(i, i+1) = \frac{A_{i+1}}{1/A_i + A_{i+1}} & i \geq 0 \\ p(i, i-1) = \frac{1/A_i}{1/A_i + A_{i+1}} & i \geq 0 \\ p(-1, 0) = 1 \end{cases}$$

Similarly we denote $\tilde{P}^\omega_i, \tilde{\mathbb{P}}_i, \tilde{E}^\omega_i, \tilde{\mathbb{E}}_i$ respectively the quenched and annealed probability/expectation for such process starting from i, and for any $n \in \mathbb{H}$, define the following stopping times

$$\tilde{\tau}_n = \inf\{k \geq 0,\ \tilde{\eta}_k = n\}, \quad \tilde{\tau}^*_n = \inf\{k \geq 1,\ \tilde{\eta}_k = n\}.$$

Let $F_1, F_2 > 0$ be two expressions which can depend on any variable, but in particular on n. If there exists $f : \mathbb{N} \to \mathbb{R}^+$ with $\lim_{n\to\infty} \frac{1}{n} \log f(n) = 0$ such that $F_1 f(n) \geq F_2$, then we denote $F_1 \gtrsim F_2$ (F_1 greater than F_2 up to polynomial constant). If $F_1 \gtrsim F_2$ and $F_1 \lesssim F_2$, then we write $F_1 \simeq F_2$.

Recall that A is Inverse Gaussian distributed with parameter $(1, c^2)$, define the rate function associated to $\log A$ by

$$I(x) = \sup_{t\in\mathbb{R}} \{tx - \log \mathbf{E}(A^t)\}, \tag{9}$$

also define

$$t^* = \sup\{t \in \mathbb{R},\ \mathbf{E}(A^t) q_1 \leq 1\}. \tag{10}$$

Lemma 2. *For any $z > 0$ and $0 < z_1 < 1$, we have, for any $0 < a < 1$*

$$\tilde{\mathbb{P}}_0(\tilde{\tau}_n \wedge \tilde{\tau}_{-1} > m | A_0 \in [a, \frac{1}{a}]) \gtrsim \exp\{-n \left(z_1 I(\frac{z}{2z_1}) + (1 - z_1) I(\frac{-z}{2(1-z_1)}) \right)\}$$

where $m \in \mathbb{N}$ is such that $n = \lfloor \frac{\log m}{z} \rfloor$.

Lemma 3. *Denote*

$$L' = \sup_{z>0,\ 0<z_1<1} \{ \frac{\log q_1}{z} - \frac{z_1}{z} I(\frac{z}{2z_1}) - \frac{1-z_1}{z} I(\frac{-z}{2(1-z_1)}) \},$$

we have $L' = -t^* + \frac{1}{2}$.

Lemma 4. *Define, for $i \in \mathbb{H}$ and any stopping time τ, $\tilde{G}^\tau(i,i) = \tilde{E}_i^\omega(\sum_{k=0}^\tau \mathbb{1}_{\tilde{\eta}_k = i})$. Let $0 \leq Y_1 < Y_2 < y < Y_3$ be points on the half line, we have, for any $0 \leq \lambda \leq 1$,*

$$\widetilde{P}_{Y_1}^\omega(\tilde{\tau}_y < \tilde{\tau}_{Y_1 - 1}) \widetilde{G}^{\tilde{\tau}_{Y_1} \wedge \tilde{\tau}_{Y_3}}(y, y) \leq \widetilde{E}_{Y_1}^\omega[\tilde{\tau}_{Y_1-1} \wedge \tilde{\tau}_{Y_3}]. \tag{11}$$

$$\widetilde{E}_{Y_1}^\omega[\tilde{\tau}_{Y_1-1} \wedge \tilde{\tau}_{Y_3}]^\lambda \leq S_{\lambda, [\![Y_1, Y_2]\!]} \left(1 + A_{Y_2+1}^\lambda \left(1 + \widetilde{E}_{Y_2+1}^\omega [\tilde{\tau}_{Y_2} \wedge \tilde{\tau}_{Y_3}]^\lambda \right) \right). \tag{12}$$

where

$$S_{\lambda, [\![Y_1, Y_2]\!]} := 1 + 2A_{Y_1}^\lambda \sum_{Y_1 < z \leq Y_2} \prod_{Y_1 < u < z} A_u^{2\lambda} A_z^\lambda + A_{Y_1}^\lambda \prod_{Y_1 < u \leq Y_2} A_u^{2\lambda}.$$

Lemma 5. *If $0 \leq \lambda < (t^* - \frac{1}{2}) \wedge 1$, then there exists sufficiently small $\delta > 0$ such that for all $n_1 > 0$*

$$\mathbf{E}\Big((1 + \frac{1}{A_{n_1}^\lambda})(1 + \frac{1}{A_n}) A_0 \widetilde{E}_0^\omega [\tilde{\tau}_{-1} \wedge \tilde{\tau}_n]^\lambda \Big) \lesssim (q_1 + \delta)^{-n}.$$

4.3 Null speed regime

In this section we prove (2) of Theorem 2.

Proposition 1. *Recall the definition of t^* in* (10), *if $q_1 \mathbf{E}(A^{-1/2}) > 1$, then $1 < t^* < \frac{3}{2}$ and*

$$\limsup_n \frac{\log |\eta_n|}{\log n} \leq t^* - \frac{1}{2}.$$

In particular, if $q_1 \mathbf{E}(A^{-1/2}) > 1$, then $\mathbb{P}$-a.s., $v(\eta) = 0$; in fact,

$$|\eta_n| = n^{(t^* - 1/2) + o(1)} = o(n),\ n \to \infty.$$

Remarks 3. *Similar arguments can be carried out for the continuous time process (Z_t), i.e. if $q_1 \mathbf{E}(A^{-1/2}) > 1$, then*

$$\limsup_t \frac{\log |Z_t|}{\log t} \leq t^* - \frac{1}{2}. \tag{13}$$

Let us state an estimate on the tail distribution of the regeneration time Γ_1 under $S(\cdot)$:

Lemma 6.

$$S(\Gamma_1 > n) \gtrsim n^{-t^* + \frac{1}{2}} \tag{14}$$

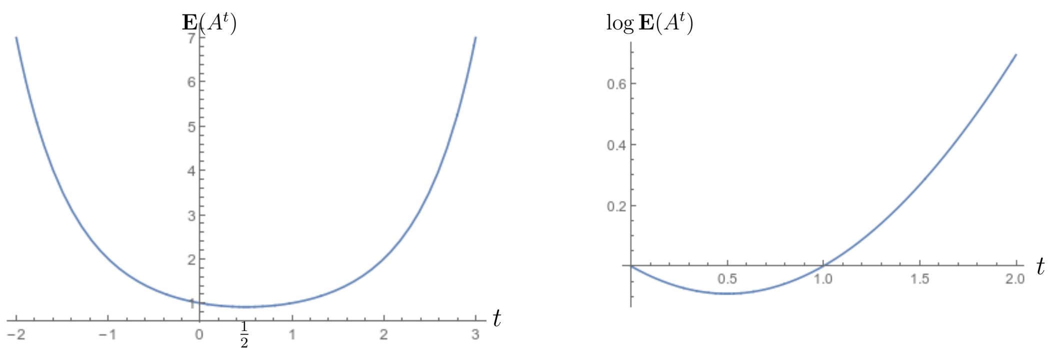

Figure 1: The function $t \mapsto \mathbf{E}(A^t)$ and $t \mapsto \log \mathbf{E}(A^t)$ for $c = 1$.

With the help of the above lemma, we prove Proposition 1.

Proof of Proposition 1. Note that $t \mapsto \mathbf{E}(A^t)$ is a convex function, and it is symmetric w.r.t. the line $t = \frac{1}{2}$, where it takes the minimum, in particular $\mathbf{E}(A^{-1/2}) = \mathbf{E}(A^{3/2})$. As we have assumed that $q_1 \mathbf{E}(A^{-1/2}) > 1$, it follows that $t^* < \frac{3}{2}$. On the other hand, since $\mathbf{E}(A) = 1$, obviously $t^* > 1$. For any $\lambda \in (t^* - 1/2, 1)$, by Lemma 6, there exists $\varepsilon > 0$ such that

$$\begin{aligned}\mathbb{P}(\max_{2 \le k \le n} (\Gamma_k - \Gamma_{k-1}) \le n^{1/\lambda}) &= S(\Gamma_1 \le n^{1/\lambda})^{n-1} \\ &\le (1 - n^{-1+\varepsilon})^{n-1} \lesssim \exp(-n^{\varepsilon}).\end{aligned}$$

Therefore,

$$\sum_{n \ge 2} \mathbb{P}(\max_{2 \le k \le n} (\Gamma_k - \Gamma_{k-1}) \le n^{1/\lambda}) < \infty.$$

By Borel-Cantelli lemma, $\mathbb{P}$-a.s., for all n large enough,

$$\Gamma_n \ge \max_{2 \le k \le n} (\Gamma_k - \Gamma_{k-1}) \ge n^{1/\lambda}.$$

It follows that $\mathbb{P}$-a.s., $\liminf_n \frac{\log \Gamma_n}{\log n} \ge \frac{1}{\lambda}$. As $\liminf_n \frac{\log \tau_n}{\log n} \ge \liminf_n \frac{\log \Gamma_n}{\log n}$ (see (3.1) in [1]), we have

$$\limsup_n \frac{\log |\eta_n|}{\log n} \le \lambda \xrightarrow[\text{decreasing}]{} t^* - \frac{1}{2} < 1, \mathbb{P}\text{-a.s.}$$

□

It remains to prove Lemma 6. In fact, when q_1 is large, it is more likely that there will be some long branch constituting vertices of degree two on the GW tree, especially starting from the root. These branches will slow down the process and entail zero velocity. The following lemma gives a comparison between the tail distribution of the regeneration time Γ_1 and the probability that the process wanders on these branches (which is a one dimensional random walk in random environment, that is, $(\tilde{\eta}_n)$).

Lemma 7. *For any $m \ge 1,\ 0 < a < 1$, we have*

$$S(\Gamma_1 > m) \ge c_5 \sum_{n=1}^{\infty} q_1^n \tilde{\mathbb{P}}_0(\tilde{\tau}_{-1} \wedge \tilde{\tau}_n > m | A_0 \in [a, \frac{1}{a}]).$$

Now we prove Lemma 6 with the help of Lemma 7 and some results on the one dimensional RWRE.

Proof of Lemma 6. By Lemma 2, one sees that for $z > 0,\ 0 < z_1 < 1$ and m such that $n = \lfloor \frac{\log m}{z} \rfloor$,

$$\tilde{\mathbb{P}}_0(\tilde{\tau}_n \wedge \tilde{\tau}_{-1} > m | A_0 \in [a, \frac{1}{a}]) \gtrsim \exp(-n \left(z_1 I(\frac{z}{2z_1}) + (1-z_1) I(\frac{-z}{2(1-z_1)}) \right))$$

where we recall that $I(x) = \sup_{t\in\mathbb{R}}\{tx - \log \mathbf{E}(A^t)\}$. For large m, by Lemma 7, then Lemma 2,

$$\begin{aligned} S(\Gamma_1 > m) &\geq c_5 \max_{n: n = \lfloor \frac{\log m}{z} \rfloor} q_1^n \tilde{\mathbb{P}}_0(\tilde{\tau}_n \wedge \tilde{\tau}_{-1} > m | A_0 \in [a, \frac{1}{a}]) \\ &\gtrsim \max_{n: n = \lfloor \frac{\log m}{z} \rfloor} q_1^n \exp(-n \left(z_1 I(\frac{z}{2z_1}) + (1-z_1) I(\frac{-z}{2(1-z_1)}) \right)) \\ &\gtrsim \sup_{z>0, z_1\in(0,1)} \exp\{-\frac{\log m}{z} \left(z_1 I(\frac{z}{2z_1}) + (1-z_1) I(\frac{-z}{2(1-z_1)}) - \log q_1 \right)\}. \end{aligned}$$

It follows from Lemma 3 that

$$S(\Gamma_1 > m) \gtrsim m^{-t^*+1/2}.$$

□

It remains to prove the comparison Lemma 7. We define, for $x \neq \overleftarrow{\rho}$,

$$\tau_x = \inf\{n \geq 0;\ \eta_n = x\},\ \tau_x^* = \inf\{n > 0;\ \eta_n = x\},\ \beta(x) = P_x^{w,T}(T_{\overleftarrow{x}} = \infty)$$

Note that for any $x \in T$, $\beta(x)$ depends only on the sub-tree T_x rooted at x and the environment $\{A_y(\omega); y \in T_x\}$, let us denote β a generic r.v. distributed as $\beta(\rho)$, by transient assumption, $\beta > 0$ a.s. and $\mathbb{E}(\beta) > 0$.

Moreover, by Markov property,

$$\begin{aligned} \beta(x) &= \sum_{y: \overleftarrow{y} = x} p(x,y)[P_y^{\omega,T}(\tau_x = \infty) + P_y^{\omega,T}(\tau_x < \infty)\beta(x)] \\ &= \sum_{y: \overleftarrow{y} = x} p(x,y)[\beta(y) + (1-\beta(y))\beta(x)]. \end{aligned}$$

Note that $\beta(x) > 0$, $\mathbb{P}$-a.s. hence,

$$\frac{1}{\beta(x)} = 1 + \frac{1}{A_x \sum_{y: \overleftarrow{y} = x} A_y \beta(y)}. \tag{15}$$

In particular, $\beta(x)$ is increasing as a function of A_x.

Proof of Lemma 7. For any vertex x, let $h(x)$ be the first descendant of x such that $d(h(x)) \geq 3$. Let $k_0 = \inf\{k \geq 2:\ q_k > 0\}$. According to the definition of Γ_1, one observes that when $\eta_1 \neq \overleftarrow{\rho}$,

$$\Gamma_1 \geq \tau_\rho^* \wedge \tau_{h(X_1)}.$$

In fact, we are going to consider the following events

$$\begin{aligned} E_0 &= \{d(\rho) = k_0 + 1,\ A_\rho \geq a,\ A_{\rho_i} \in [a, \frac{1}{a}], \forall 1 \leq i \leq k_0\} \text{ where } \rho_i \text{ are children of } \rho, \\ E_1 &= E_0 \cap \{\eta_1 \neq \overleftarrow{\rho}, m < \tau_\rho^* < \tau_{h(\eta_1)}, \eta_{\tau_\rho^*+1} \notin \{\overleftarrow{\rho}, \eta_1\}\} \cap \{\eta_n \neq \rho; \forall n \geq \tau_\rho^* + 1\}, \\ E_2 &= E_0 \cap \{\eta_1 \neq \overleftarrow{\rho}, m < \tau_{h(\eta_1)} < \tau_\rho^*\} \cap \{\eta_n \neq \overleftarrow{h(\eta_1)}, \forall n \geq \tau_{h(\eta_1)} + 1\}. \end{aligned}$$

As $\Gamma_1 \geq \tau_\rho^* \wedge \tau_{h(\eta_1)}$, we have $E_1 \cup E_2 \subset E_0 \cap \{D(\rho) = \infty, \Gamma_1 > m\}$ and $E_1 \cap E_2 = \emptyset$. So,

$$\mathbb{P}(E_0 \cap \{D(\rho) = \infty, \Gamma_1 > m\}) \geq \mathbb{P}(E_1) + \mathbb{P}(E_2).$$

For E_1, by strong Markov property at τ_ρ^* and weak Markov property at time 1,

$$\begin{aligned}
P_\rho^{\omega,T}(E_1) &= \mathbb{1}_{E_0} P_\rho^{\omega,T}(\{\eta_1 \neq \overleftarrow{\rho}, m < \tau_\rho^* < \tau_{h(\eta_1)}, \eta_{\tau_\rho^*+1} \notin \{\overleftarrow{\rho}, \eta_1\}\} \cap \{\eta_n \neq \rho; \forall n \geq \tau_\rho^* + 1\}) \\
&= \mathbb{1}_{E_0} \sum_{i=1}^{k_0} p(\rho, \rho_i) P_{\rho_i}^{\omega,T}(m - 1 < \tau_\rho < \tau_{h(\rho_i)}) \sum_{j \neq i} p(\rho, \rho_j) \beta(\rho_j).
\end{aligned}$$

Given E_0, $p(\rho, \rho_i) \geq \frac{a^2}{k_0+1} =: c_6$. So,

$$P_\rho^{\omega,T}(E_1) \geq c_6 \mathbb{1}_{E_0} \sum_{i=1}^{k_0} P_{\rho_i}^{\omega,T}(m - 1 < \tau_\rho < \tau_{h(\rho_i)}) \sum_{j \neq i} p(\rho, \rho_j) \beta(\rho_j),$$

Conditionally on $\{d(\rho), A_\rho, A_{\rho_i}, 1 \leq i \leq d(\rho) - 1\}$, the independence of the environment implies that

$$\begin{aligned}
&\mathbb{P}\Big(E_1 \Big| d(\rho), A_\rho, A_{\rho_i}, 1 \leq i \leq d(\rho) - 1\Big) \\
&\qquad\qquad \geq c_6 \mathbb{1}_{E_0} \sum_{i=1}^{k_0} \mathbb{P}_{\rho_i}(m - 1 < \tau_\rho < \tau_{h(\rho_i)}) \sum_{j \neq i} p(\rho, \rho_j) \mathbf{E}_{\mathbf{Q}}[\beta(\rho_j) | A_{\rho_j}],
\end{aligned}$$

where, for each $j \neq i$, $p(\rho, \rho_j)$ and $\mathbf{E}_{\mathbf{Q}}[\beta(\rho_j)|A_{\rho_j}]$ are increasing functions of A_{ρ_j}. By FKG inequality, and the fact that $\mathbb{E}(\beta(\rho)) > 0$ and $\sum_{j \neq i} p(\rho, \rho_j) \geq \frac{a^2(k_0-1)}{1+k_0} > 0$ on E_0,

$$\begin{aligned}
\mathbb{P}(E_1) &\geq c_6 \mathbb{E}\Big(\mathbb{1}_{E_0} \sum_{i=1}^{k_0} P_{\rho_i}^{\omega,T}(m - 1 < \tau_\rho < \tau_{h(\rho_i)}) \sum_{j \neq i} p(\rho, \rho_j)\Big) \times \mathbb{E}(\beta(\rho)) \\
&\geq c_7 \mathbb{E}\Big(\mathbb{1}_{E_0} \sum_{i=1}^{k_0} P_{\rho_i}^{\omega,T}(m - 1 < \tau_\rho < \tau_{h(\rho_i)})\Big).
\end{aligned} \tag{16}$$

Similarly for E_2, by Markov property,

$$\begin{aligned}
P_\rho^{\omega,T}(E_2) &= \mathbb{1}_{E_0} P_\rho^{\omega,T}(\{\eta_1 \neq \overleftarrow{\rho}, m < \tau_{h(\eta_1)} < \tau_\rho^*\} \cap \{\eta_n \neq \overleftarrow{h(\eta_1)}; \forall n \geq \tau_{h(\eta_1)} + 1\}) \\
&= \mathbb{1}_{E_0} \sum_{i=1}^{k_0} p(\rho, \rho_i) P_{\rho_i}^{\omega,T}(m - 1 < \tau_{h(\rho_i)} < \tau_\rho) \beta(h(\rho_i)) \\
&\geq c_6 \mathbb{1}_{E_0} \sum_{i=1}^{k_0} P_{\rho_i}^{\omega,T}(m - 1 < \tau_{h(\rho_i)} < \tau_\rho) \beta(h(\rho_i)).
\end{aligned}$$

Again $P_{\rho_i}^{\omega,T}(m - 1 < \tau_{h(\rho_i)} < \tau_\rho)$ and $\beta(h(\rho_j))$ are both increasing on $A_{h(\rho_i)}$. FKG inequality entails

$$\begin{aligned}
\mathbb{P}(E_2) &\geq c_6 \mathbb{E}\Big(\mathbb{1}_{E_0} \sum_{i=1}^{k_0} P_{\rho_i}^{\omega,T}(m - 1 < \tau_{h(\rho_i)} < \tau_\rho)\Big) \times \mathbb{E}(\beta(\rho)) \\
&= c_8 \mathbb{E}(\mathbb{1}_{E_0} \sum_{i=1}^{k_0} P_{\rho_i}^{\omega,T}(m - 1 < \tau_{h(\rho_i)} < \tau_\rho)),
\end{aligned} \tag{17}$$

with $c_8 := c_6\mathbb{E}(\beta(\rho)) > 0$. Combining (16) with (17) yields that

$$\begin{aligned}\mathbb{P}(E_1) + \mathbb{P}(E_2) &\geq c_9\mathbb{E}\Big(\mathbb{1}_{E_0}\sum_{i=1}^{k_0} P_{\rho_i}^{\omega,T}(\tau_\rho \wedge \tau_{h(\rho_i)} > m-1)\Big) \\ &\geq c_9 K_0 \mathbf{Q}(E_0)\mathbb{P}\Big(\tau_{\overleftarrow{\rho}} \wedge \tau_{h(\rho)} > m-1 | A_\rho \in [a, \frac{1}{a}]\Big) \\ &\geq c_{10}\mathbb{P}\Big(\tau_{\overleftarrow{\rho}} \wedge \tau_{h(\rho)} > m-1 | A_\rho \in [a, \frac{1}{a}]\Big). \end{aligned} \tag{18}$$

Let us go back to $S(\Gamma_1 > m)$. As $\mathbb{P}(d(\rho) \geq 3, D(\rho) = \infty) > 0$, recall that

$$\begin{aligned} S(\Gamma_1 > m) &= \mathbb{P}(\Gamma_1 > m | d(\rho) \geq 3, D(\rho) = \infty) \\ &\geq \mathbb{P}(E_0 \cap \{D(\rho) = \infty, \Gamma_1 > m\}) \\ &\geq \mathbb{P}(E_1) + \mathbb{P}(E_2). \end{aligned}$$

by (18), taking $c_5 = c_{10}$, we have

$$\begin{aligned} S(\Gamma_1 > m) &\geq c_5\mathbb{P}\Big(\tau_{\overleftarrow{\rho}} \wedge \tau_{h(\rho)} > m-1 | A_\rho \in [a, \frac{1}{a}]\Big) \\ &= c_5 \sum_{n=1}^{\infty} q_1^n \tilde{\mathbb{P}}_0(\tilde{\tau}_{-1} \wedge \tilde{\tau}_n > m-1 | A_0 \in [a, \frac{1}{a}]). \end{aligned}$$

□

4.4 Positive speed on big tree and asymptotic of $|Z_t|$ on small tree

This subsection is devoted to the proof of the following propositions, firstly when the tree is big (i.e. q_1 small), the RWRE has positive speed; when the tree is small (q_1 large), we can compute exactly the asymptotic behavior of $|\eta_n|$ and $|Z_t|$.

Proposition 2. *If $q_1\mathbf{E}(A^{-1/2}) < 1$, then*

$$v(\eta) > 0 \text{ and } v(Z) > 0. \tag{19}$$

As a consequence, also $v(Y) > 0$.

Proposition 3. *Assume that $q_1\mathbf{E}(A^{-1/2}) > 1$, we have $\mathbb{P}$-a.s.*

$$\lim_{n\to\infty} \frac{\log|\eta_n|}{\log n} = \lim_{t\to\infty} \frac{\log|Z_t|}{\log t} = t^* - 1/2 \in (1/2, 1) \tag{20}$$

where $t^* = \sup\{t \in \mathbb{R},\ \mathbf{E}(A^t)q_1 \leq 1\}$.

Let us give some definitions and heuristics before proving these propositions, write, for $n \geq 0$,

$$\tau_n(\eta) = \inf\{k \geq 0;\ |\eta_k| = n\} \text{ and } \tau_n(Z) = \inf\{t \geq 0;\ |Z_t| = n\}$$

the hitting times of the n-th generation for η and Z respectively. As a consequence of the law of large numbers, $\mathbb{P}$-a.s.,

$$\lim_{n\to\infty} \frac{\tau_n(\eta)}{n} = \frac{1}{v(\eta)} \text{ and } \lim_{n\to\infty} \frac{\tau_n(Z)}{n} = \frac{1}{v(Z)}.$$

The study of the speed is reduced to the study of $\tau_n(\eta)$ and $\tau_n(Z)$. For any $x \in T$, $n \geq -1$, let N_x and N_n denote the time spent by the walk η at x and at the n-th generation respectively:

$$N(x) = \sum_{k\geq 0} \mathbb{1}_{\eta_k = x}, \quad N_n = \sum_{|x|=n} N(x),$$

observe that

$$\tau_n(\eta) \leq \sum_{k=-1}^{n} N_k, \quad E^{\omega,T}[\tau_n(Z)|\eta] \leq \sum_{x:-1\leq|x|\leq n} N_x \frac{A_x}{1+A_xB_x},$$

where $B_x := \sum_{y:\overleftarrow{y}=x} A_y$.

In what follows, we actually study N_n for large n to show that $\liminf_n \frac{\sum_{k=-1}^n N_k}{n} < \infty$, $\mathbb{P}$-a.s. The heuristics is the following. Fix some n_0, K_0 (to choose later), pick some vertex y at the n-th generation, if y roughly lies in a subtree of height n_0 with more than K_0 leaves, then the random walk will immediately go down, thus $\mathbb{E}(N_y)$ will be small c.f. Figure 2 left. Otherwise, we seek a down going path $\hat{y}, \dots, y, \dots, \check{y}$ such that every vertex in this path does not branch much except for the two ends, and we need these two ends have more than K_0 descendants after n_0 generations. In such configuration, we can compare the random walk to the one dimensional one, and once the walker reaches one of the ends, it immediately leaves our path $\hat{y}, \dots, \check{y}$ c.f. Figure 2 right.

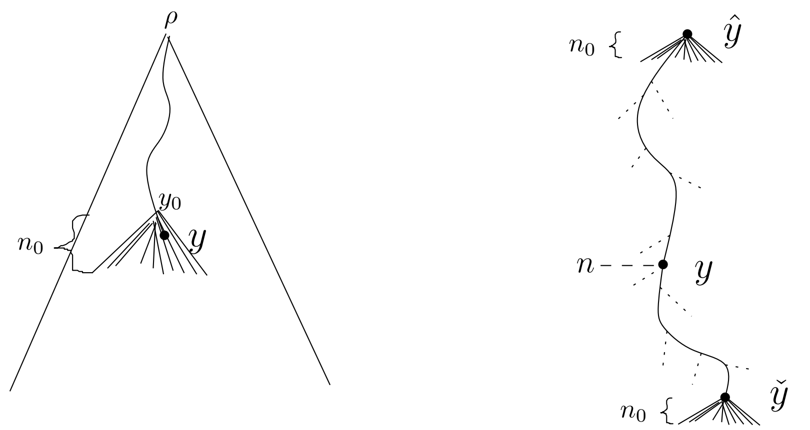

Figure 2: Two cases to bound $\mathbb{E}(N_y)$.

If the root have more than K_0 descendants after n_0 generations, then we can always find $\hat{y}$. Otherwise, we need to take n large and use the Galton Watson structure. To handle this issue, let us introduce the following notations. For the GW tree T, let Z_n^T be the number of vertices at the n-th generation. By Lemma 4.1 of [1], we have for any $K_0 \geq 1$,

$$\mathbb{E}_{\mathsf{GW}}(Z_n^T \mathbb{1}_{Z_n^T \leq K_0}) \leq K_0 n^{K_0} q_1^{n-K_0}.$$

Let $r \in (q_1, 1)$ be some real we choose later, let

$$n_0 = n_0(K_0, r) := \inf\{n \geq 1, \ \mathbb{E}_{\mathsf{GW}}(Z_n^T \mathbb{1}_{Z_n^T \leq K_0}) \leq r^n\},$$

which is thus a finite integer. In fact, K_0 will be chosen according to Corollary 3. Define

$$Z^T(u, n) = |\{x \in T; \ u < x, |x| = |u| + n\}|.$$

Let T_{n_0} be a tree induced from T in the following way: starting from the root ρ, y is a child of x in T_{n_0} if $x < y$ and $|y| = |x| + n_0$. Define a subtree $\mathcal{W}$ of T_{n_0} by

$$\mathcal{W} = \{x \in T_{n_0} : \ \forall u \in T_{n_0}, u < x \Rightarrow Z^T(u, n_0) \leq K_0\}.$$

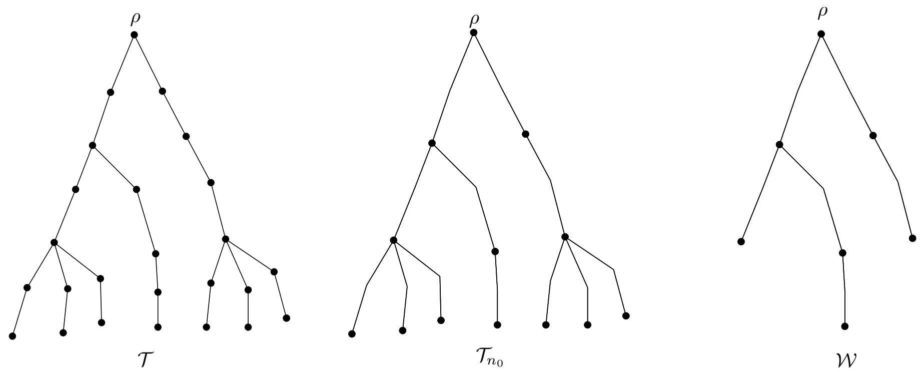

Figure 3: An example in the case $K_0 = n_0 = 2$.

Let W_k be the population of the k-th generation of $\mathcal{W}$, $\mathcal{W}$ is a sub critical Galton Watson tree of mean offspring $\mathbb{E}_{\mathsf{GW}}(Z^T_{n_0}\mathbb{1}_{Z^T_{n_0}\le K_0}) \le r^{n_0}$; in particular, for any $k \ge 0$, $\mathbb{E}_{\mathsf{GW}}(W_k) \le r^{kn_0}$.

For any $y \in T$, let y_0 be the youngest ancestor of y in T_{n_0}. For $n \ge n_0$, let $j = \lfloor \frac{n}{n_0} \rfloor \ge 1$ so that $jn_0 \le n < (j+1)n_0$. Define

$$N_{n,1} = \sum_{|y|=n} N(y)\mathbb{1}_{Z^T(y_0,n_0)>K_0}, \quad N^*_{n,1} = \sum_{|y|=n} N(y)\frac{A_y}{1+A_yB_y}\mathbb{1}_{Z^T(y_0,n_0)>K_0} \tag{21}$$

$$N_{n,2} = \sum_{|y|=n} N(y)\mathbb{1}_{Z^T(y_0,n_0)\le K_0,\ y_0\notin\mathcal{W}}, \quad N^*_{n,2} = \sum_{|y|=n} N(y)\frac{A_y}{1+A_yB_y}\mathbb{1}_{Z^T(y_0,n_0)\le K_0,\ y_0\notin\mathcal{W}} \tag{22}$$

Lemma 8. *There exist $r \in (q_1, 1)$ and $K_0 > 0$, such that, with the definitions of $n_0, N_{n,1}, N^*_{n,1}$ above, for some constant $L > 0$, for any $n \ge n_0$*

$$\mathbb{E}(N_{n,1}) \le L, \quad \mathbb{E}(N^*_{n,1}) \le L. \tag{23}$$

Lemma 9. *With the same assumption as in Lemma 8, if $0 < \lambda < 1 \wedge (t^* - 1/2)$ where t^* is define in* (10)*, then*

$$\mathbb{E}(N^\lambda_{n,2}) \le L, \quad \mathbb{E}((N^*_{n,2})^\lambda) \le L. \tag{24}$$

We are prepared to prove Proposition 2 and Proposition 3.

Proof of Proposition 2. Since $q_1\mathbf{E}(A^{-1/2}) < 1$, $t^* > 3/2$. We choose $\lambda = 1$. As $\mathcal{W}$ is finite a.s., if $\chi = (\mathrm{height}(\mathcal{W}) + 1)n_0$ (where for a finite tree T, $\mathrm{height}(T) := \max_{x\in T}|x|$), then

$$\text{for all } n \ge \chi,\ N_n \le N_{n,1} + N_{n,2}.$$

By Lemma 8, 9, for any $n \ge n_0$,

$$\mathbb{E}(N_n, n \ge \chi) \le 2L.$$

Thus,

$$\liminf_{n\to\infty} \mathbb{E}\Big[\frac{\sum_{i=\chi}^n N_n}{n}\Big] \le 2L.$$

By Fatou's lemma, a.s.

$$\liminf_{n\to\infty} \frac{\sum_{k=-1}^n N_k}{n} = \liminf_{n\to\infty} \frac{\sum_{k=\chi}^n N_k}{n} < \infty.$$

Therefore,

$$\frac{1}{v(\eta)} = \liminf_{n\to\infty} \frac{\tau_n}{n} \le \liminf_{n\to\infty} \frac{\sum_{k=-1}^{n} N_k}{n} < \infty.$$

This implies that $v(\eta) > 0$.

The case for Z_t can be treated in a similar manner with N_n^* instead of N_n. Finally, to prove $v(Y) > 0$, it is enough to recall $Z_{D(t)} = Y_t$ where $D(t) = \sum_x (l_x(t)^2 + 2cl_x(t))$ and note that

$$\frac{D(t)}{t} = \frac{\sum_x (l_x(t)^2 + 2cl_x(t))}{\sum_x l_x(t)} \ge 2c > 0.$$

It follows that

$$v(Y) = \lim_{t\to\infty} \frac{|Y_t|}{t} = \lim_{t\to\infty} \frac{|Z_{D(t)}|}{t} \ge v(Z) \liminf_{t\to\infty} \frac{D(t)}{t} \ge 2cv(Z).$$

□

Proof of Proposition 3. If $q_1\mathbf{E}(A^{-1/2}) \ge 1$, $\lambda < t^* - 1/2 \le 1$. Let $N_i(Z)$ be the time spent at the i-th generation by (Z_t). Let $\Gamma_k(Z)$ be the regenerative times corresponding to $(Z_t)_{t\ge 0}$. Let $u(n)$ be the unique integer such that $\Gamma_{u(n)} \le \tau_n(Z) < \Gamma_{u(n)+1}$. Then,

$$\begin{aligned}\frac{\Gamma_{u(n)}(Z)^\lambda}{n} \le \frac{\sum_{k\le u(n)} (\Gamma_k(Z) - \Gamma_{k-1}(Z))^\lambda}{n} &= \frac{\sum_{k\le u(n)} (\sum_{i=|Z_{\Gamma_{k-1}(Z)}|}^{i=|Z_{\Gamma_k(Z)}|-1} N_i(Z))^\lambda}{n} \\ &\le \frac{\sum_{i\le n} N_i(Z)^\lambda}{n}.\end{aligned}$$

Taking limit yields that

$$\liminf_{n\to\infty} \frac{\Gamma_{u(n)}(Z)^\lambda}{n} \le \liminf_{n\to\infty} \frac{\sum_{k\le u(n)} (\Gamma_k(Z) - \Gamma_{k-1}(Z))^\lambda}{n} \le \liminf_{n\to\infty} \frac{\sum_{i=\chi}^{n} N_i(Z)^\lambda}{n}.$$

Applying Jensen's inequality then Lemma 9 implies that

$$\mathbb{E}[N_n(Z)^\lambda; n \ge \chi] \le \mathbb{E}[\mathbf{E}[N_n(Z); n \ge \chi|\eta]^\lambda] \le \mathbb{E}[(N_n^*)^\lambda, n \ge \chi] \le 2L.$$

It follows from Fatou's lemma that

$$\liminf_{n\to\infty} \frac{\Gamma_{u(n)}(Z)^\lambda}{n} \le \liminf_{n\to\infty} \frac{\sum_{k\le u(n)} (\Gamma_k(Z) - \Gamma_{k-1}(Z))^\lambda}{n} < \infty.$$

By law of large numbers,

$$\lim_{n\to\infty} \frac{n}{u(n)} = \mathbb{E}_S[|Z_{\Gamma_1(Z)}|] < \infty, \text{ and } \lim_{n\to\infty} \frac{\sum_{k\le n} (\Gamma_k(Z) - \Gamma_{k-1}(Z))^\lambda}{n} = \mathbb{E}_S[\Gamma_1(Z)^\lambda].$$

Therefore there exists a constant $C \in (0, \infty)$ such that

$$\liminf_{n\to\infty} \frac{\Gamma_n(Z)^\lambda}{n} < C.$$

Note that $|Z_t| \ge \#\{k : \Gamma_k(Z) < t\}$. So we get $|Z_t| \ge t^\lambda / C$ for all sufficiently large t. We hence deduce that

$$\liminf_{t\to\infty} \frac{\log |Z_t|}{\log t} \ge \lambda.$$

Letting $\lambda \uparrow t^* - 1/2$ yields

$$\liminf_{n\to\infty} \frac{\log|Z_t|}{\log t} \geq t^* - 1/2. \tag{25}$$

The result follows by Remark 3. Similar arguments can be applied to $\lim_{n\to\infty} \frac{\log|\eta_n|}{\log n}$. $\square$

It remains to show the main Lemmas 8,9. Let us first state some preliminary results. As the walk is transient, the support of the random walk should be slim. This is formulated in the following lemma:

Lemma 10. *There exists a constant $c_{11} > 0$ such that for any $n \geq 1$, $\mathbb{E}(\sum_{|x|=n} \mathbb{1}_{\tau_x<\infty}) \leq c_{11}$.*

The following lemma shows that, the escape probability is relatively large. In fact, we cannot show that $\mathbb{E}(\frac{1}{\beta(\rho)}) < \infty$ for all $q_1 > 0$, since the GW tree branches anyway, there will be a large copies of independent sub-trees, we show $\mathbb{E}(\frac{1}{\sum_{i=1}^K \beta_i}) < \infty$ instead.

Lemma 11. *Consider i.i.d. copies of GW trees $T^{(i)}$ rooted at $\rho^{(i)}$ with independent environment $\omega^{(i)}$, for each $T^{(i)}$, define $\beta_i = P^{\omega^{(i)},T^{(i)}}_{\rho^{(i)}}(\tau_{\overleftarrow{\rho^{(i)}}} = \infty)$. There exists an integer $K = K(q_1, c) \geq 1$ such that*

$$\mathbb{E}(\frac{1}{\sum_{i=1}^K \beta_i}) \leq c_{12} < \infty \quad \textit{and} \quad \mathbb{E}(\frac{1}{\sum_{i=1}^K A_{\rho^{(i)}}\beta_i}) < c_{12} < \infty.$$

Moreover, if $q_1\xi_2 < 1$, then $\mathbb{E}(\dfrac{1}{\beta(\rho)}) \leq c_{12} < \infty$ and $\mathbb{E}(\frac{1}{A_\rho\beta(\rho)}) < c_{12} < \infty$.

Remarks 4. *In fact, if $q_1\mathbf{E}(A^{-2}) < 1$, a proof similar to Proposition 2.3 of [1] shows that η has positive speed, in particular, the VRJP on any regular tree (except $\mathbb{Z}$) admits positive speed.*

Corollary 3. *There exists $K_0 \geq K$, such that*

$$\mathbb{E}(\frac{1}{\sum_1^{K_0} A^2_{\rho^{(i)}}\beta_i^2}) < c_{13} < \infty.$$

The proof of Lemma 10, 11 and Corollary 3 will be postponed to the Appendix B, let us state the consequence of these preliminary results. Recall that Z_n^T is the population at generation n, and that for any $x \in T$, τ_x is the first hitting time, τ_x^* the first return time to x. For $u, v \in T$ write $u < v$ if u is an ancestor of v and define

$$p_1(u, v) = P_u^{\omega,T}(\tau_{\overleftarrow{u}} = \infty, \tau_u^* = \infty, \tau_v = \infty)$$

Lemma 12. *For any $n \geq 2$ and $k \in \{1, 2\}$, consider K_0 as in Corollary 3, we have*

$$\mathbb{E}\Big(\mathbb{1}_{Z_n^T>K_0} \sum_{|u|=n} \frac{1}{p_1(\rho, u)^k}\Big) < c_{14}^n < \infty.$$

In addition,

$$\mathbb{E}\Big(\mathbb{1}_{Z_n^T>K_0} \sum_{|u|=n} \frac{1}{p_1(\rho, u)^k}\Big|A_\rho\Big) < c_{14}^n(1 + \frac{1}{A_\rho}). \tag{26}$$

Proof of Lemma 12. Fix $n \geq 2$, let $\Upsilon_0 := \inf\{l \geq 1;\ Z_l > K_0\}$, then $\{Z_n^T > K_0\} = \{\Upsilon_0 \leq n\}$. For any $u \in T$ such that $|u| \geq \Upsilon_0$, let U be its ancestor at the Υ_0-th generation. By Markov property,

$$\begin{aligned} p_1(\rho, u) &\geq \sum_{|y|=\Upsilon_0-1} P_\rho^{\omega,T}(\tau_y < \tau_\rho^*) P_y^{\omega,T}(\tau_{\overleftarrow{y}} = \infty, \tau_U = \infty) \\ &\geq \sum_{|y|=\Upsilon_0-1} \prod_{i=0}^{\Upsilon_0-2} p(y_i, y_{i+1}) P_y^{\omega,T}(\tau_{\overleftarrow{y}} = \infty, \tau_U = \infty) \end{aligned} \tag{27}$$

where $\{y_0(=\rho), y_1, \ldots, y_{\Upsilon_0-1}(=y)\}$ is the unique path connecting ρ and y. Note that if $\overleftarrow{U} = y$, then

$$P_y^{\omega,T}(\tau_{\overleftarrow{y}} = \infty, \tau_U = \infty) = \sum_{z:\overleftarrow{z}=y, z\neq U} p(y,z)\beta(z) + \sum_{z:\overleftarrow{z}=y, z\neq U} p(y,z)(1-\beta(z)) P_y^{\omega,T}(\tau_{\overleftarrow{y}} = \infty, \tau_U = \infty).$$

Otherwise

$$P_y^{\omega,T}(\tau_{\overleftarrow{y}} = \infty, \tau_U = \infty) = \sum_{z:\overleftarrow{z}=y} p(y,z)\beta(z) + \sum_{z:\overleftarrow{z}=y} p(y,z)(1-\beta(z)) P_y^{\omega,T}(\tau_{\overleftarrow{y}} = \infty, \tau_U = \infty)$$

It follows that in both cases,

$$\begin{aligned} P_y^{\omega,T}(\tau_{\overleftarrow{y}} = \infty, \tau_U = \infty) &= \frac{\sum_{z:\overleftarrow{z}=y} \mathbb{1}_{z\neq U} p(y,z)\beta(z)}{p(y, \overleftarrow{y}) + p(y,U) + \sum_{z:\overleftarrow{z}=y} \mathbb{1}_{z\neq U} p(y,z)\beta(z)} \\ &\geq \frac{\sum_{z:\overleftarrow{z}=y} \mathbb{1}_{z\neq U} A_y A_z \beta(z)}{1 + A_y A_U + \sum_{z:\overleftarrow{z}=y} \mathbb{1}_{z\neq U} A_y A_z \beta(z)} \\ &\geq \frac{A_y}{1+A_y} \frac{1}{1+A_U} \frac{\sum_{z:\overleftarrow{z}=y} \mathbb{1}_{z\neq U} A_z \beta(z)}{1 + \sum_{z:\overleftarrow{z}=y} \mathbb{1}_{z\neq U} A_z \beta(z)} \end{aligned}$$

Plugging it into (27) yields that

$$\begin{aligned} p_1(\rho, u) &\geq \sum_{|y|=\Upsilon_0-1} \prod_{i=0}^{\Upsilon_0-2} p(y_i, y_{i+1}) \frac{A_y}{1+A_y} \frac{1}{1+A_U} \frac{\sum_{z:\overleftarrow{z}=y} \mathbb{1}_{z\neq U} A_z \beta(z)}{1 + \sum_{z:\overleftarrow{z}=y} \mathbb{1}_{z\neq U} A_z \beta(z)} \\ &\geq \frac{1}{1+A_U} \min_{|y|=\Upsilon_0-1} \left(\prod_{i=0}^{\Upsilon_0-2} p(y_i, y_{i+1}) \frac{A_y}{1+A_y} \right) \cdot \frac{\sum_{z:|z|=\Upsilon_0, z\neq U} A_z \beta(z)}{1 + \sum_{z:|z|=\Upsilon_0, z\neq U} A_z \beta(z)} \end{aligned}$$

Thus, for $k \in \{1, 2\}$,

$$\frac{1}{p_1(\rho,u)^k} \leq (1+A_U)^k \frac{1}{\min_{|y|=\Upsilon_0-1} \left(\prod_{i=0}^{\Upsilon_0-2} p(y_i, y_{i+1}) \frac{A_y}{1+A_y} \right)^k} \Big(1 + \frac{1}{\sum_{z:|z|=\Upsilon_0, z\neq U} A_z \beta(z)}\Big)^k.$$

Given the tree T, by integrating w.r.t. $\mathbf{P}(d\omega)$, we have

$$\begin{aligned} \mathbb{1}_{n\geq \Upsilon_0} \sum_{|u|=n} \mathbb{E}^T \Big(\frac{1}{p_1(\rho,u)^k} \Big) &\leq \mathbb{E}^T \left(\frac{1}{\min_{|y|=\Upsilon_0-1} \left(\prod_{i=0}^{\Upsilon_0-2} p(y_i, y_{i+1}) \frac{A_y}{1+A_y} \right)^k} \right) \\ &\quad \times \sum_{|U|=\Upsilon_0} Z^T(U, n-\Upsilon_0) \mathbb{E}^T[(1+A_U)^k] \mathbb{E}^T \left(\Big(1 + \frac{1}{\sum_{z:|z|=\Upsilon_0, z\neq U} A_z \beta(z)} \Big)^k \right) \end{aligned}$$

It follows from Lemma 11 for $k = 1$ or Corollary 3 for $k = 2$ that

$$\begin{aligned}
&\mathbf{E}_{\mathbf{Q}}\left(\mathbb{1}_{n\geq\Upsilon_0}\sum_{|u|=n}\frac{1}{p_1(\rho,u)^k}\middle|\Upsilon_0, Z_l; 0\leq l\leq\Upsilon_0\right)\\
\leq& c_{15}\mathbb{1}_{n\geq\Upsilon_0}\mathbb{E}^T\left(\frac{1}{\min_{|y|=\Upsilon_0-1}\left(\prod_{i=0}^{\Upsilon_0-2}p(y_i,y_{i+1})\frac{A_y}{1+A_y}\right)^k}\right)\times\sum_{|U|=\Upsilon_0}\mathbf{E}[(1+A)^k]b^{n-\Upsilon_0}\\
\leq& c_{16}\mathbb{1}_{n\geq\Upsilon_0}\sum_{|y|=\Upsilon_0-1}\mathbb{E}^T\left[\left(\prod_{i=0}^{\Upsilon_0-2}\frac{(1+A_{y_i})(1+B_{y_i})}{A_{y_i}A_{y_{i+1}}}\frac{1+A_y}{A_y}\right)^k\right]\sum_{|U|=\Upsilon_0}b^{n-\Upsilon_0}.
\end{aligned}$$

By independence of $A_x, x\in T$, we see that

$$\mathbb{E}^T\left[\left(\prod_{i=0}^{\Upsilon_0-2}\frac{(1+A_{y_i})(1+B_{y_i})}{A_{y_i}A_{y_{i+1}}}\frac{1+A_y}{A_y}\right)^k\right]\leq c_{17}^{\Upsilon_0-1},$$

with $c_{17}\in(1,\infty)$. Consequently,

$$\begin{aligned}
\mathbf{E}_{\mathbf{Q}}\left(\mathbb{1}_{n\geq\Upsilon_0}\sum_{|u|=n}\frac{1}{p_1(\rho,u)^k}\right)\leq&\ \mathbf{E}_{\mathbf{Q}}\Big(c_{16}\mathbb{1}_{n\geq\Upsilon_0}\sum_{|y|=\Upsilon_0-1}c_{15}^{\Upsilon_0-1}\sum_{|U|=\Upsilon_0}b^{n-\Upsilon_0}\Big)\\
\leq& c_{16}K_0\mathbf{E}_{\mathbf{Q}}\Big(\mathbb{1}_{n\geq\Upsilon_0}c_{15}^{n-1}Z_n^T\Big)\\
\leq& c_{18}(c_{17}b)^n<\infty.
\end{aligned}$$

(26) follows in the same way. $\square$

Proof of Lemma 8. We only bound $\mathbb{E}(N_{n,1})$, the argument for $\mathbb{E}(N^*_{n,1})$ is similar. For any $y\in T$ at the n-th generation such that $Z^T(y_0,n_0)>K_0$, let Y be the youngest ancestor of y such that $Z^T(Y,n_0)>K_0$. Clearly, $y_0\leq Y\leq y$. So,

$$N_{n,1}=\sum_{|y|=n}N(y)\mathbb{1}_{Z^T(y_0,n_0)>K_0}\leq\sum_{|y|=n}N(y)\mathbb{1}_{y_0\leq Y\leq y}.$$

Taking expectation w.r.t. $E_\rho^{\omega,T}$ implies that

$$E^{\omega,T}(N_{n,1})\leq\sum_{|y|=n}E^{\omega,T}(N(y))\mathbb{1}_{y_0\leq Y\leq y}=\sum_{|y|=n}P^{\omega,T}(\tau_y<\infty)E_y^{\omega,T}(N(y))\mathbb{1}_{y_0\leq Y\leq y}.$$

Applying the Markov property at τ_Y to $E_y^{\omega,T}(N(y))$, we have

$$E_y^{\omega,T}(N(y))=G^{\tau_Y}(y,y)+P_y^{\omega,T}(\tau_Y<\infty)P_Y^{\omega,T}(\tau_y<\infty)E_y^{\omega,T}(N(y))$$

where (write $\{(\tau_Y\wedge\infty)>\tau_y^*\}=\{\tau_y^*<\infty$ and $\tau_y^*<\tau_Y\}$ for short)

$$G^{\tau_Y}(y,y)=E_y^{\omega,T}(\sum_{k=0}^{\tau_Y}\mathbb{1}_{\eta_k=y})=\frac{1}{1-P_y^{\omega,T}((\tau_Y\wedge\infty)>\tau_y^*)}.$$

Hence

$$E_y^{\omega,T}(N(y)) = \frac{G^{\tau_Y}(y,y)}{1 - P_Y^{\omega,T}(\tau_y < \infty)P_y^{\omega,T}(\tau_Y < \infty)}$$
$$\le \frac{G^{\tau_Y}(y,y)}{1 - P_Y^{\omega,T}(\tau_Y^* < \infty)} = \frac{G^{\tau_Y}(y,y)}{P_Y^{\omega,T}(\tau_Y^* = \infty)}.$$

We bound $G^{\tau_Y}(y,y)$ first. As $P_y^{\omega,T}((\tau_Y \wedge \infty) > \tau_y^*) \le \sum_{z:\overleftarrow{z}=y} p(y,z) + p(y,\overleftarrow{y})P_{\overleftarrow{y}}^{\omega,T}(\tau_y < (\tau_Y \wedge \infty))$,

$$1 - P_y^{\omega,T}((\tau_Y \wedge \infty) > \tau_y^*) \ge p(y,\overleftarrow{y})\Big(1 - P_{\overleftarrow{y}}^{\omega,T}(\tau_y < \tau_Y)\Big).$$

By Lemma 4.4 of [1] and (38), the right hand side of the above inequality is larger than

$$p(y,\overleftarrow{y})\Big(1 - \tilde{P}_{\overleftarrow{y}}^{\omega,T}(\tilde{\tau}_y < \tilde{\tau}_Y)\Big) = \frac{1}{1 + A_y B_y}\frac{1}{1 + A_y \sum_{Y<z<y} A_z \prod_{z<u<y} A_u^2}.$$

where we identify $\tilde{P}_{\overleftarrow{y}}^{\omega,T}$ to the probability of $(\tilde{\eta}_n)$ on the segment $[\![Y, y]\!]$. Therefore,

$$G^{\tau_Y}(y,y) \le \Big(1 + A_y \sum_{Y<z<y} A_z \prod_{z<u<y} A_u^2\Big)(1 + A_y B_y) =: V_{y,Y}.$$

Consequently,

$$E^{\omega,T}(N(y))\mathbb{1}_{Z^T(y_0,n_0)>K_0} \le P^{\omega,T}(\tau_Y < \infty)\frac{V_{y,Y}}{P_Y^{\omega,T}(\tau_Y^* = \infty)}\mathbb{1}_{Z^T(Y,n_0)>K_0,\ y_0 \le Y \le y}.$$

Summing over all possibilities of Y yields that (recall that $j = \lfloor \frac{n}{n_0} \rfloor$)

$$E^{\omega,T}(N_{n,1}) \le \sum_{l=jn_0}^{n} \sum_{|Y|=l} P^{\omega,T}(\tau_Y < \infty)\frac{\sum_{|y|=n, Y\le y} V_{y,Y}}{P_Y^{\omega,T}(\tau_Y^* = \infty)}\mathbb{1}_{Z^T(Y,n_0)>K_0}$$
$$\le \sum_{l=jn_0}^{n} \sum_{|Y|=l} P^{\omega,T}(\tau_{\overleftarrow{Y}} < \infty)\frac{\sum_{|y|=n, Y\le y} V_{y,Y}}{P_Y^{\omega,T}(\tau_Y^* = \infty, \tau_{\overleftarrow{Y}} = \infty)}\mathbb{1}_{Z^T(Y,n_0)>K_0},$$

where the last inequality holds because $P^{\omega,T}(\tau_Y < \infty) \le P^{\omega,T}(\tau_{\overleftarrow{Y}} < \infty)$ and $P_Y^{\omega,T}(\tau_Y^* = \infty) \ge P_Y^{\omega,T}(\tau_Y^* = \infty, \tau_{\overleftarrow{Y}} = \infty)$. Summing over the value of $\overleftarrow{Y}$ yields that

$$E^{\omega,T}(N_{n,1}) \le \sum_{l=jn_0-1}^{n-1} \sum_{|x|=l} P^{\omega,T}(\tau_x < \infty) \sum_{Y:\overleftarrow{Y}=x} \frac{\sum_{|y|=n, Y\le y} V_{y,Y}}{P_Y(\tau_Y^* = \infty, \tau_{\overleftarrow{Y}} = \infty)}\mathbb{1}_{Z^T(Y,n_0)>K_0}.$$

As conditionally on T, $P^{\omega,T}(\tau_x < \infty)$ and $\sum_{Y:\overleftarrow{Y}=x} \frac{\sum_{|y|=n,Y\le y} V_{y,Y}}{P_Y(\tau_Y^*=\infty,\tau_{\overleftarrow{Y}}=\infty)}\mathbb{1}_{d(Y,n_0)>K_0}$ are independent,

$$\mathbb{E}(N_{n,1}) \le \mathbb{E}\left(\sum_{l=jn_0-1}^{n-1} \sum_{|x|=l} \mathbb{E}^T(P^{\omega,T}(\tau_x < \infty))\mathbb{E}^T\Big(\sum_{Y:\overleftarrow{Y}=x} \frac{\sum_{|y|=n,Y\le y} V_{y,Y}}{P_Y^{\omega,T}(\tau_Y^* = \infty, \tau_{\overleftarrow{Y}} = \infty)}\mathbb{1}_{Z^T(Y,n_0)>K_0}\Big)\right)$$
$$= \sum_{l=jn_0-1}^{n-1} \mathbb{E}(\sum_{|x|=l} \mathbb{1}_{\tau_x<\infty})\mathbb{E}\Big(\sum_{|Y|=1} \frac{\sum_{|y|=n-l,Y\le y} V_{y,Y}}{P_Y^{\omega,T}(\tau_Y^* = \infty, \tau_{\overleftarrow{Y}} = \infty)}\mathbb{1}_{Z^T(Y,n_0)>K_0}\Big)$$

Note that for any $|Y|=1$, $\frac{\sum_{|y|=n-l, Y\le y} V_{y,Y}}{P_Y^{\omega,T}(\tau_Y^*=\infty, \tau_{\overleftarrow{Y}}=\infty)}\mathbb{1}_{Z^T(Y,n_0)>K_0}$ are i.i.d. By Lemma 10,

$$\mathbb{E}(N_{n,1}) \le bc_{11}\sum_{l=jn_0-1}^{n-1}\mathcal{A}_{n-l} \tag{28}$$

where

$$\mathcal{A}_{n-l} = \mathbb{E}\Big(\frac{\sum_{|y|=n-l-1} V_{y,\rho}}{P^{\omega,T}(\tau_\rho^*=\infty, \tau_{\overleftarrow{\rho}}=\infty)}\mathbb{1}_{Z^T(\rho,n_0)>K_0}\Big).$$

By Cauchy-Schwartz inequality,

$$\mathcal{A}_{n-l} \le \mathbb{E}\Big[\Big(\sum_{|y|=n-l-1} V_{y,\rho}\Big)^2\Big]\mathbb{E}\Big[\frac{\mathbb{1}_{Z^T(\rho,n_0)>K_0}}{P^{\omega,T}(\tau_\rho^*=\infty, \tau_{\overleftarrow{\rho}}=\infty)^2}\Big]$$

Recall that Z_n^T denote the number of vertices at the n-th generation of the tree T, using Lemma 12 then Applying again Cauchy-Schwartz inequality to $\Big(\sum_{|y|=n-l-1} V_{y,\rho}\Big)^2$ implies that

$$\begin{aligned}\mathcal{A}_{n-l} &\le c_{14}^{n_0}\mathbb{E}\Big(Z_{n-l-1}^T\sum_{|y|=n-l-1} V_{y,\rho}^2\Big)\\ &\le c_{19}\mathbf{E}_{GW}[c_{20}^{n-l-1}\Big(Z_{n-l-1}^T\Big)^2],\end{aligned}$$

where the second inequality follows from $\mathbb{E}^T[V_{y,\rho}]\le c_{20}^{|y|}$. Plugging it into (28) implies that

$$\mathbb{E}(N_{n,1}) \le bc_{11}c_{19}\sum_{l=jn_0-1}^{n-1}\mathbf{E}_{GW}[c_{20}^{n-l-1}\Big(Z_{n-l-1}^T\Big)^2] \le c_{21}\sum_{k=0}^{n_0}c_{20}^k\mathbf{E}_{GW}\Big[(Z_k^T)^2\Big] \le c_{22},$$

since $\mathbf{E}_{GW}[(Z_1^T)^2]<\infty$. Analoguesly, for $N_{n,1}^*$ we get that

$$E^{\omega,T}(N_{n,1}^*) \le \sum_{l=jn_0-1}^{n-1}\sum_{|x|=l}P^{\omega,T}(\tau_x<\infty)\sum_{Y:\overleftarrow{Y}=x}\frac{\sum_{|y|=n, Y\le y} V_{y,Y}\frac{A_y}{1+A_yB_y}}{P_Y^{\omega,T}(\tau_Y^*=\infty, \tau_{\overleftarrow{Y}}=\infty)}\mathbb{1}_{Z^T(Y,n_0)>K_0}.$$

And recounting on the same arguments gives a finite upper bound for $\mathbb{E}[N_{n,1}^*]$. □

Proof of Lemma 9. Again we only give the proof for $\mathbb{E}(N_{n,2}^\lambda)$. For $y\in T$, as $Z^T(y_0,n_0)\le K_0$ and $y_0\notin\mathcal{W}$, we can find the youngest ancestor Y_1 of y in T_{n_0} such that $Z^T(Y_1,n_0)>K_0$, automatically $Y_1<y_0$. Let Y_2 be the youngest descendant of Y_1 in T_{n_0} such that it is an ancestor of y. Let Y_3 be the youngest descendant of y in T_{n_0} such that $Z^T(Y_3,n_0)>K_0$.
For any $0<\lambda\le 1$,

$$\begin{aligned}E^{\omega.T}[N_{n,2}^\lambda] &\le E^{\omega,T}\Big[\sum_{|y|=n}N(y)^\lambda\mathbb{1}_{Z^T(y_0,n_0)\le K_0,\ y_0\notin\mathcal{W}}\Big]\\ &\le \sum_{|y|=n}\mathbb{1}_{Z^T(y_0,n_0)\le K_0,\ y_0\notin\mathcal{W}}P^{\omega,T}(\tau_y<\infty)\Big(E_y^{\omega,T}[N(y)]\Big)^\lambda.\end{aligned} \tag{29}$$

In what follows, we identify $\widetilde{P}^\omega$ with the distribution of a one-dimensional random walk $\widetilde{\eta}$ on the path $[\![\overleftarrow{Y_1},\ Y_3]\!]$. Let us state the following lemmas which will be used in (29).

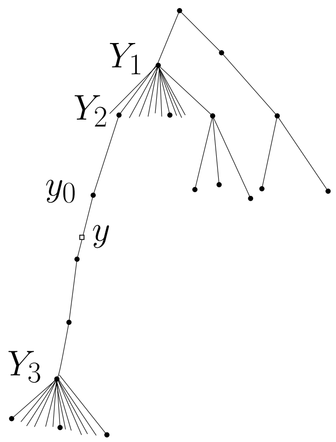

Figure 4: An example of Y_1, Y_2, Y_3.

Lemma 13. *For any $y \in T$ such that $Y_1 < Y_2 < y < Y_3$, let y^* be the unique child of y which is also ancestor of Y_3. Then,*

$$\Big(E_y^{\omega,T}[N(y)]\Big)^\lambda \leq \Big(\frac{1+A_yB_y}{1+A_yA_{y^*}}\Big)^\lambda \widetilde{G}^{\tilde{\tau}_{Y_1}\wedge\tilde{\tau}_{Y_3}}(y,y)^\lambda \frac{2}{p_1(Y_1,Y_2)P_{Y_3}^{\omega,T}(\tau^*_{Y_3}=\infty, \tau_{\overleftarrow{Y_3}}=\infty)}. \tag{30}$$

where $\widetilde{G}^{\tilde{\tau}_{Y_1}\wedge\tilde{\tau}_{Y_3}}(y,y) = \widetilde{E}_y^\omega\Big(\sum_{k=0}^{\tilde{\tau}_{Y_1}\wedge\tilde{\tau}_{Y_3}} \mathbb{1}_{\widetilde{\eta}_k=y}\Big)$ is the Green function associated with $(\widetilde{\eta}_n)$.

Lemma 14.

$$P^{\omega,T}(\tau_y<\infty) \leq P^{\omega,T}(\tau_{Y_1}<\infty)\widetilde{P}^\omega_{Y_1}(\tilde{\tau}_y < \tilde{\tau}_{Y_1-1})^\lambda \frac{1}{p_1(Y_1,Y_2)}. \tag{31}$$

The proofs of Lemmas 13 and 14 can be found in section 5.2 of [1] with slight modifications, so we feel free to omit them (see (5.10) and (5.11) therein). Now plugging (30) and (31) into (29) yields that

$$E^{\omega,T}(N^\lambda_{n,2}) \leq \sum_{|y|=n} \frac{2P^{\omega,T}(\tau_{Y_1}<\infty)}{p_1(Y_1,Y_2)^2 P_{Y_3}^{\omega,T}(\tau^*_{Y_3}=\infty, \tau_{\overleftarrow{Y_3}}=\infty)} \bigg(\frac{1+A_yB_y}{1+A_yA_{y^*}} \widetilde{P}^\omega_{Y_1}(\tilde{\tau}_y < \tilde{\tau}_{Y_1-1}) \widetilde{G}^{\tilde{\tau}_{Y_1}\wedge\tilde{\tau}_{Y_3}}(y,y)\bigg)^\lambda.$$

By Lemma 4, one sees that

$$\begin{aligned}
E^{\omega,T}(N^\lambda_{n,2}) &\leq \sum_{|y|=n} \frac{2P^{\omega,T}(\tau_{Y_1}<\infty)}{p_1(Y_1,Y_2)^2 P_{Y_3}^{\omega,T}(\tau^*_{Y_3}=\infty, \tau_{\overleftarrow{Y_3}}=\infty)} \bigg(\frac{1+A_yB_y}{1+A_yA_{y^*}} \widetilde{E}^\omega_{Y_1}[\tilde{\tau}_{\overleftarrow{Y_1}} \wedge \tilde{\tau}_{Y_3}]\bigg)^\lambda \\
&\leq \sum_{|y|=n} \frac{2P^{\omega,T}(\tau_{Y_1}<\infty)}{p_1(Y_1,Y_2)^2 P_{Y_3}^{\omega,T}(\tau^*_{Y_3}=\infty, \tau_{\overleftarrow{Y_3}}=\infty)} \Big(\frac{1+A_yB_y}{1+A_yA_{y^*}}\Big)^\lambda S_{\lambda,[\![Y_1,Y_2]\!]}\bigg(1+A^\lambda_{Y_2^*}\Big(1+\widetilde{E}^\omega_{Y_2^*}[\tilde{\tau}_{Y_2}\wedge\tilde{\tau}_{Y_3}]^\lambda\Big)\bigg)
\end{aligned}$$

where Y_2^* is the children of Y_2 along $[\![Y_2, Y_3]\!]$. Decompose the sum over $|y|=n$ by

$$\sum_{|y|=n} = \sum_{y:|y|=n, Y_1=\rho} + \sum_{l=1}^{(j-1)} \sum_{|x|=ln_0-1} \sum_{y:\overleftarrow{Y_1}=x, |y|=n}.$$

We get that

$$E^{\omega,T}(N_{n,2}^{\lambda}) \le \sum_{|y|=n, Y_1=\rho} \frac{2S_{\lambda,[\![\rho,Y_2]\!]}}{p_1(\rho,Y_2)^2 P^{\omega,T}_{Y_3}(\tau^*_{Y_3}=\infty, \tau_{\overleftarrow{Y_3}}=\infty)}\Theta_\lambda(Y_2,y,Y_3)$$
$$+\sum_{l=1}^{j-1}\sum_{|x|=ln_0-1}\sum_{|y|=n,\overleftarrow{Y_1}=x} \frac{2P^{\omega,T}(\tau_{Y_1}<\infty)S_{\lambda,[\![Y_1,Y_2]\!]}}{p_1(Y_1,Y_2)^2 P^{\omega,T}_{Y_3}(\tau^*_{Y_3}=\infty, \tau_{\overleftarrow{Y_3}}=\infty)}\Theta_\lambda(Y_2,y,Y_3),$$

where

$$\Theta_\lambda(Y_2,y,Y_3) := \left(\frac{1+A_yB_y}{1+A_yA_{y^*}}\right)^{\lambda}\left(1+A^{\lambda}_{Y_2^*}\left(1+\widetilde{E}^{\omega}_{Y_2^*}[\tilde\tau_{Y_2}\wedge\tilde\tau_{Y_3}]^{\lambda}\right)\right).$$

Given the GW tree T, note that $S_{\lambda,[\![Y_1,Y_2]\!]}\in\sigma\{A_z; Y_1\le z\le Y_2\}$, $p_1(\rho,Y_2)\in\sigma\{A_u : u\in(T\setminus T_{Y_2})\cup\{Y_2\}\}$, $P^{\omega,T}_{Y_3}(\tau^*_{Y_3}=\infty,\tau_{\overleftarrow{Y_3}}=\infty)\in\sigma\{A_u; u\in T_{Y_3}\}$ and $\Theta_\lambda(Y_2,y,Y_3)\in\sigma\{A_u; Y_2<u\le Y_3\}$. Therefore,

$$\mathbb{E}^T[N^\lambda_{n,2}] \le \sum_{|y|=n,Y_1=\rho}\mathbb{E}^T\Big[\frac{2S_{\lambda,[\![\rho,Y_2]\!]}}{p_1(\rho,Y_2)^2}\Big]\mathbb{E}^T\Big[\frac{\Theta_\lambda(Y_2,y,Y_3)\mathbb{1}_{Z^T(Y_3,n_0)>K_0}}{P^{\omega,T}_{Y_3}(\tau^*_{Y_3}=\infty, \tau_{\overleftarrow{Y_3}}=\infty)}\Big]$$
$$+\sum_{l=1}^{j-1}\sum_{|x|=ln_0-1}\sum_{|y|=n,\overleftarrow{Y_1}=x}\mathbb{E}^T\Big[\frac{2P^{\omega,T}(\tau_{Y_1}<\infty)S_{\lambda,[\![Y_1,Y_2]\!]}}{p_1(Y_1,Y_2)^2}\Big]\mathbb{E}^T\Big[\frac{\Theta_\lambda(Y_2,y,Y_3)\mathbb{1}_{Z^T(Y_3,n_0)>K_0}}{P^{\omega,T}_{Y_3}(\tau^*_{Y_3}=\infty, \tau_{\overleftarrow{Y_3}}=\infty)}\Big]. \tag{32}$$

Observe that

$$P^{\omega,T}_{Y_3}(\tau^*_{Y_3}=\infty,\tau_{\overleftarrow{Y_3}}=\infty)\ge p_1(Y_3,u)\mathbb{1}_{Y_3<u,|u|=|Y_3|+n_0}.$$

$$\mathbb{E}^T\Big[\frac{\Theta_\lambda(Y_2,y,Y_3)\mathbb{1}_{Z^T(Y_3,n_0)>K_0}}{P^{\omega,T}_{Y_3}(\tau^*_{Y_3}=\infty, \tau_{\overleftarrow{Y_3}}=\infty)}\Big|A_u, Y_2<u\le Y_3\Big] = \Theta_\lambda(Y_2,y,Y_3)\mathbb{E}\Big[\frac{\mathbb{1}_{Z^T(Y_3,n_0)>K_0}}{P^{\omega,T}_{Y_3}(\tau^*_{Y_3}=\infty, \tau_{\overleftarrow{Y_3}}=\infty)}\Big|A_{Y_3}\Big]$$
$$\le\Theta_\lambda(Y_2,y,Y_3)\mathbb{E}\Big[\mathbb{1}_{Z^T(Y_3,n_0)>K_0}\sum_{u:Y_3<u,|u|=|Y_3|+n_0}\frac{1}{p_1(Y_3,u)}\Big|A_{Y_3}\Big].$$

Applying Lemma 12 to the subtree rooted at Y_3 implies that

$$\mathbb{E}^T\Big[\frac{\Theta_\lambda(Y_2,y,Y_3)\mathbb{1}_{Z^T(Y_3,n_0)>K_0}}{P^{\omega,T}_{Y_3}(\tau^*_{Y_3}=\infty, \tau_{\overleftarrow{Y_3}}=\infty)}\Big]\le c_{23}\mathbb{E}^T\Big[(1+\frac{1}{A_{Y_3}})\Theta_\lambda(Y_2,y,Y_3)\Big].$$

Plugging it into (32) implies that

$$\mathbb{E}^T[N^\lambda_{n,2}]\le\Delta_1(n)+\Delta_2(n),$$

where

$$\Delta_1(n) := 2c_{23}\sum_{|y|=n,Y_1=\rho}\mathbb{E}^T\Big[\frac{S_{\lambda,[\![\rho,Y_2]\!]}}{p_1(\rho,Y_2)^2}\Big]\mathbb{E}^T\Big[(1+\frac{1}{A_{Y_3}})\Theta_\lambda(Y_2,y,Y_3)\Big] \tag{33}$$
$$\Delta_2(n) := 2c_{23}\sum_{l=1}^{j-1}\sum_{|x|=ln_0-1}\sum_{|y|=n,\overleftarrow{Y_1}=x}\mathbb{E}^T\Big[\frac{P^{\omega,T}(\tau_{Y_1}<\infty)S_{\lambda,[\![Y_1,Y_2]\!]}}{p_1(Y_1,Y_2)^2}\Big]\mathbb{E}^T\Big[(1+\frac{1}{A_{Y_3}})\Theta_\lambda(Y_2,y,Y_3)\Big]. \tag{34}$$

So,

$$\mathbb{E}[N^{\lambda}_{n,2}] \leq \mathbf{E}_{\mathbf{Q}}[\Delta_1(n) + \Delta_2(n)]. \tag{35}$$

We firstly bound $\Delta_1(n)$, note that (since $\lambda \leq 1$)

$$\Big(\frac{1 + A_y B_y}{1 + A_y A_{y^*}}\Big)^{\lambda} \leq \Big(1 + \frac{\sum_{z:\overleftarrow{z}=y, z\neq y^*} A_z}{A_{y^*}}\Big)^{\lambda} \leq 1 + \frac{\sum_{z:\overleftarrow{z}=y, z\neq y^*} A_z^{\lambda}}{A_{y^*}^{\lambda}},$$

with $\sum_{z:\overleftarrow{z}=y,z\neq y^*} 1 \leq K_0$. If $|Y_2| = mn_0 < n$, $|Y_3| = (m+k)n_0 > n$, by Markov property and the fact that $\{A_z, \overleftarrow{z} = y, z \neq y^*\}$ is independent of $\{A_z, z \in [\![Y_2, Y_3]\!] := [\![-1, kn_0 - 1]\!]\}$,

$$\begin{aligned}
&\mathbb{E}^T\Big[(1 + \frac{1}{A_{Y_3}})\Theta_{\lambda}(Y_2, y, Y_3)\Big] \\
&\leq \mathbb{E}^T\left[(1 + \frac{1}{A_{kn_0-1}})(1 + \frac{\sum_{z:\overleftarrow{z}=y,z\neq y^*} A_z^{\lambda}}{A^{\lambda}_{n-mn_0}})(1 + A_0^{\lambda}(1 + \tilde{E}^{\omega}_0(\tilde{\tau}_{-1} \wedge \tilde{\tau}_{kn_0-1})^{\lambda}))\right] \\
&\leq c_{24} + c_{24}\mathbf{E}\Big((1 + \frac{1}{A_{kn_0-1}})(1 + \frac{1}{A^{\lambda}_{n-mn_0}})A_0^{\lambda}\widetilde{E}^{\omega}_0[\tilde{\tau}_{-1} \wedge \tilde{\tau}_{kn_0-1}]^{\lambda}\Big).
\end{aligned}$$

Now apply Lemma 5, we have

$$\mathbb{E}^T\Big[(1 + \frac{1}{A_{Y_3}})\Theta_{\lambda}(Y_2, y, Y_3)\Big] \leq c_{25}(q_1 + \delta)^{-|Y_3|+|Y_2|+1}. \tag{36}$$

Applying Cauchy-Schwartz inequality to $\mathbb{E}^T\Big[\frac{S_{\lambda,[\![\rho,Y_2]\!]}}{p_1(\rho,Y_2)^2}\Big]$ yields

$$\begin{aligned}
\Delta_1(n) &\leq c_{23} \sum_{|y|=n, Y_1=\rho} 2\Big(\sqrt{\mathbb{E}^T\Big[S^2_{\lambda,[\![\rho,Y_2]\!]}\Big]\mathbb{E}^T\Big[\frac{1}{p_1(\rho, Y_2)^4}\Big]}\Big)\mathbb{E}^T\Big[(1 + \frac{1}{A_{Y_3}})\Theta_{\lambda}(Y_2, y, Y_3)\Big] \\
&\leq c_{26} \sum_{|y|=n, Y_1=\rho} \sqrt{\mathbb{E}^T\Big[\frac{1}{p_1(\rho, Y_2)^4}\Big]}\mathbb{E}^T\Big[(1 + \frac{1}{A_{Y_3}})\Theta_{\lambda}(Y_2, y, Y_3)\Big],
\end{aligned}$$

where the last inequality holds because $\mathbb{E}^T\Big[S^2_{\lambda,[\![\rho,Y_2]\!]}\Big] \leq c_{27}(n_0) < \infty$. By (36),

$$\begin{aligned}
\Delta_1(n) &\leq c_{28} \sum_{|y|=n, Y_1=\rho} \mathbb{E}^T\Big[\frac{1}{p_1(\rho, Y_2)^4}\Big](q_1 + \delta)^{-|Y_3|+|Y_2|+1} \\
&= c_{28}\mathbb{E}^T\Big[\sum_{|u|=n_0} \mathbb{1}_{Z^T_{n_0} > K_0} \frac{1}{p_1(\rho, u)^4}\Big] \sum_{y:|y|=n, Y_2=u} (q_1 + \delta)^{-|Y_3|+n_0+1}
\end{aligned}$$

Observe that

$$\sum_{y:|y|=n, Y_2=u} (q_1 + \delta)^{-|Y_3|+n_0+1} \leq \sum_{z:|z|>n, z\in\mathcal{W}(T_u)} (q_1 + \delta)^{-|z|+n_0+1}.$$

Hence,

$$\Delta_1(n) \leq c_{28}\mathbb{E}^T\Big[\sum_{|u|=n_0} \mathbb{1}_{Z^T_{n_0} > K_0} \frac{1}{p_1(\rho, u)^4}\Big] \sum_{z:|z|>n, z\in\mathcal{W}(T_u)} (q_1 + \delta)^{-|z|+n_0+1}.$$

Taking expectation under $GW(dT)$ implies that

$$\mathbf{E}_{\mathbf{Q}}[\Delta_1(n)] \le c_{28}\mathbb{E}\Big[\sum_{|u|=n_0} \mathbb{1}_{Z^T_{n_0}>K_0}\frac{1}{p_1(\rho,u)^4}\Big]\mathbf{E}_{\mathbf{Q}}\Big[\sum_{z:|z|>n-n_0, z\in\mathcal{W}} (q_1+\delta)^{-|z|+1}\Big],$$

which by Lemma 12 is bounded by

$$c_{29}\mathbf{E}_{\mathbf{Q}}\Big[\sum_{z:|z|>n-n_0, z\in\mathcal{W}} (q_1+\delta)^{-|z|+1}\Big] = c_{29}\sum_{l>n/n_0-1}\mathbf{E}_{\mathbf{Q}}\Big[\sum_{|z|=ln_0, z\in\mathcal{W}} (q_1+\delta)^{-|z|+1}\Big].$$

Recall that $\mathcal{W}$ is a GW tree of mean $\mathbf{E}[Z_{n_0}; Z_{n_0}\le K_0]\le r^{n_0}$. We can choose r to be $q_1+\delta/2$ so that

$$\sum_{l\ge 1}\mathbf{E}_{\mathbf{Q}}\Big[\sum_{|z|=ln_0, z\in\mathcal{W}} (q_1+\delta)^{-|z|+1}\Big] \le \sum_{l\ge 1}(q_1+\delta)^{-ln_0+1}r^{ln_0} < c_{30}\gamma^{l_0},$$

where $\gamma := (\frac{q_1+\delta/2}{q_1+\delta})^{n_0}<1$ and $l_0 := \lceil\frac{n}{n_0}\rceil - 1 = j-1$. As a result, for any $n>n_0$,

$$\mathbf{E}_{\mathbf{Q}}[\Delta_1(n)] \le c_{31}\gamma^{l_0} < \infty. \tag{37}$$

Turn to $\Delta_2(n)$. As $P^{\omega,T}(\tau_{Y_1}<\infty)\le P^{\omega,T}(\tau_{\overleftarrow{Y_1}}<\infty)$, one sees that

$$\Delta_2(n) \le 2c_{23}\sum_{l=1}^{j-1}\sum_{|x|=ln_0-1}\sum_{|y|=n,\overleftarrow{Y_1}=x}\mathbb{E}^T\Big[\frac{P^{\omega,T}(\tau_x<\infty)S_{\lambda,[\![Y_1,Y_2]\!]}}{p_1(Y_1,Y_2)^2}\Big]\mathbb{E}^T\Big[(1+\frac{1}{A_{Y_3}})\Theta_\lambda(Y_2,y,Y_3)\Big],$$

which equals to

$$\sum_{l=1}^{j-1}\sum_{|x|=ln_0-1}\sum_{z:\overleftarrow{z}=x}\mathbb{P}^T(\tau_x<\infty)2c_{23}\sum_{|y|=n,Y_1=z}\mathbb{E}^T\Big[\frac{S_{\lambda,[\![Y_1,Y_2]\!]}}{p_1(Y_1,Y_2)^2}\Big]\mathbb{E}^T\Big[(1+\frac{1}{A_{Y_3}})\Theta_\lambda(Y_2,y,Y_3)\Big],$$

as $P^{\omega,T}(\tau_x<\infty)$ and $\frac{S_{\lambda,[\![Y_1,Y_2]\!]}}{p_1(Y_1,Y_2)^2}$ are independent under $\mathbb{P}^T$.

Note that for all $z\in T$, $2c_{23}\sum_{|y|=n,Y_1=z}\mathbb{E}^T\Big[\frac{S_{\lambda,[\![Y_1,Y_2]\!]}}{p_1(Y_1,Y_2)^2}\Big]\mathbb{E}^T\Big[(1+\frac{1}{A_{Y_3}})\Theta_\lambda(Y_2,y,Y_3)\Big]$ are i.i.d. copies of $\Delta_1(n-|z|)$. Taking expectation yields that

$$\begin{aligned}\mathbf{E}_{\mathbf{Q}}[\Delta_2(n)] &\le \sum_{l=1}^{j-1}\mathbb{E}\Big[\sum_{|x|=ln_0-1}\mathbb{1}_{\tau_x<\infty}(d(x)-1)\Big]\mathbf{E}_{\mathbf{Q}}[\Delta_1(n-ln_0)]\\ &\le bc_{31}\sum_{l=1}^{j-1}\mathbb{E}\Big[\sum_{|x|=ln_0-1}\mathbb{1}_{\tau_x<\infty}\Big]\gamma^{j-l-1},\end{aligned}$$

where the last inequality follows from (37). By Lemma 10, for any $j\ge 2$,

$$\mathbf{E}_{\mathbf{Q}}[\Delta_2(n)] \le c_{32}\sum_{l=1}^{j-1}\gamma^{j-1-l} \le c_{33} < \infty.$$

Plugging the above inequality and (37) into (35) implies that

$$\mathbb{E}[N^\lambda_{n,2}] \le \mathbf{E}_{\mathbf{Q}}[\Delta_1(n)] + \mathbf{E}_{\mathbf{Q}}[\Delta_2(n)] < \infty.$$

The estimate of $\mathbb{E}[(N^*_{n,2})^\lambda]$ follows from similar arguments. We feel free to omit it. $\square$

A Proofs of one dimensional results

Proof of Lemma 2. For any $i \geq 1$, let $S_i = -\sum_{j=1}^{i} \log(A_j A_{j-1})$ and define $S_0 = 0$. As $i \mapsto \tilde{P}_i^\omega(\tilde{\tau}_{-1} > \tilde{\tau}_n)$ is the solution to the Dirichlet problem

$$\begin{cases} \varphi(-1) = 0, \ \varphi(n) = 1 \\ \tilde{E}_i^\omega(\varphi(\tilde{\eta}_1)) = \varphi(i) \qquad i \in [\![0, n-1]\!]. \end{cases}$$

It follows that

$$\tilde{P}_i^\omega(\tilde{\tau}_{-1} > \tilde{\tau}_n) = \frac{\sum_{j=0}^{i} \exp(S_j)}{\sum_{j=0}^{n} \exp(S_j)}. \tag{38}$$

As a consequence, for any $0 \leq l \leq n$,

$$\begin{aligned} \tilde{P}_0^\omega(\tilde{\tau}_l < \tilde{\tau}_{-1}) &= \frac{1}{\sum_{j=0}^{l} \exp(S_j)} \geq \frac{\exp(-\max_{0\leq j\leq l} S_j)}{l+1} \\ \tilde{P}_{l+1}^\omega(\tilde{\tau}_n < \tilde{\tau}_l) &= \frac{\exp(S_{l+1})}{\sum_{j=l+1}^{n} \exp(S_j)} \leq \exp(-\max_{l+1\leq j\leq n}(S_j - S_{l+1})) \\ \tilde{P}_{l-1}^\omega(\tilde{\tau}_{-1} < \tilde{\tau}_l) &= \frac{\exp(S_l)}{\sum_{j=0}^{l} \exp(S_j)} \leq \exp(-\max_{0\leq j\leq l}(S_j - S_l)). \end{aligned}$$

We only need to consider n large, take $l = \lfloor z_1 n \rfloor$, note that

$$\begin{aligned} \tilde{P}_l^\omega(\tilde{\tau}_l^* > \tilde{\tau}_{-1} \wedge \tilde{\tau}_n) &= p(l, l+1)\tilde{P}_{l+1}^\omega(\tilde{\tau}_n < \tilde{\tau}_l) + p(l, l-1)\tilde{P}_{l-1}^\omega(\tilde{\tau}_{-1} < \tilde{\tau}_l) \\ &\leq \max(\tilde{P}_{l+1}^\omega(\tilde{\tau}_n < \tilde{\tau}_l), \tilde{P}_{l-1}^\omega(\tilde{\tau}_{-1} < \tilde{\tau}_l)). \end{aligned}$$

Therefore,

$$\begin{aligned} \tilde{P}_0^\omega(\tilde{\tau}_n \wedge \tilde{\tau}_{-1} > m) &\geq \tilde{P}_0^\omega(\tilde{\tau}_l < \tilde{\tau}_{-1})\tilde{P}_l^\omega(\tilde{\tau}_l^* < \tilde{\tau}_{-1} \wedge \tilde{\tau}_n)^m \\ &\geq \frac{\exp(-\max_{0\leq j\leq l} S_j)}{l+1}\left(1 - \tilde{P}_l^\omega(\tilde{\tau}_l^* \geq \tilde{\tau}_{-1} \wedge \tilde{\tau}_n)\right)^m \\ &\geq \frac{\exp(-\max_{0\leq j\leq l} S_j)}{l+1}\left(1 - \exp(-\max_{l+1\leq k\leq n}(S_k - S_{l+1}) \wedge \max_{0\leq k\leq l}(S_k - S_l))\right)^m \\ &\geq \frac{\mathbb{1}_{\max_{0\leq k\leq l} S_k \leq 0}}{l+1}(1 - e^{-zn})^m \mathbb{1}_{\max_{l+1\leq k\leq n}(S_k - S_{l+1})\geq zn}\mathbb{1}_{\max_{0\leq k\leq l}(S_k - S_l)\geq zn}. \end{aligned}$$

As $m \approx e^{zn}$, we have $(1 - e^{-zn})^m = O(1)$, taking expectation under $\mathbf{P}(\cdot|A_0 \in [a, \frac{1}{a}])$ yields

$$\begin{aligned} &\tilde{\mathbb{P}}_0(\tilde{\tau}_n \wedge \tilde{\tau}_{-1} > m | A_0 \in [a, \frac{1}{a}]) \\ &\geq \frac{c}{n}\mathbf{P}(\max_{0\leq k\leq l} S_k \leq 0, \ \max_{0\leq k\leq l}(S_k - S_l) \geq zn | A_0 \in [a, \frac{1}{a}])\mathbf{P}(\max_{l+1\leq k\leq n}(S_k - S_{l+1}) \geq zn) \\ &\geq \frac{c}{n}\mathbf{P}(\max_{0\leq k\leq l} S_k \leq 0, \ S_l \leq -zn | A_0 \in [a, \frac{1}{a}])\mathbf{P}((S_n - S_{l+1}) \geq zn). \end{aligned}$$

For $k \geq 1$, write $\mathscr{S}_k = -\sum_{i=1}^{k} \log A_i$, then as $S_k = -\log A_0 + \mathscr{S}_{k-1} + \mathscr{S}_k$,

$$\begin{aligned} &\mathbf{P}(\max_{0\leq k\leq l} S_k \leq 0, \ S_l \leq -zn | A_0 \in [a, \frac{1}{a}]) \\ &\geq \mathbf{P}(A_0 \geq 1, A_l \geq 1, \max_{1\leq k\leq l-1} \mathscr{S}_k \leq 0, \ \mathscr{S}_{l-1} \leq -\frac{zn}{2} | A_0 \in [a, \frac{1}{a}]) \\ &= \mathbf{P}(A_0 \geq 1 | A_0 \in [a, \frac{1}{a}])\mathbf{P}(A_l \geq 1)\mathbf{P}(\max_{1\leq k\leq l-1} \mathscr{S}_k \leq 0, \ \mathscr{S}_{l-1} \leq -\frac{zn}{2}) \end{aligned}$$

note that

$$\mathbf{P}(\max_{1\le k\le l-1}\mathscr{S}_k \le 0,\ \mathscr{S}_{l-1}\le -\frac{zn}{2}) \ge \frac{1}{l}\mathbf{P}(\mathscr{S}_{l-1}\le -\frac{zn}{2})$$

and

$$S_n - S_{l+1} = -\log A_{l+1} - \log A_n - 2\sum_{k=l+2}^{n-1}\log A_k.$$

Therefore,

$$\begin{aligned}
\tilde{\mathbb{P}}_0(\tilde\tau_n\wedge\tilde\tau_{-1} > m|A_0\in[a,\frac{1}{a}]) &\ge \frac{c}{n^2}\mathbf{P}(\mathscr{S}_{l-1}\le -\frac{zn}{2})\mathbf{P}(S_n - S_{l+1}\ge zn)\\
&\ge \frac{c}{n^2}\mathbf{P}(\mathscr{S}_{l-1}\le -\frac{zn}{2})\mathbf{P}(A_{l+1}\le 1)\mathbf{P}(A_n\le 1)\mathbf{P}(-\sum_{k=l+2}^{n-1}\log A_k\ge \frac{zn}{2})\\
&\ge \frac{c}{n^2}\mathbf{P}(\mathscr{S}_{l-1}\le -\frac{zn}{2})\mathbf{P}(-\sum_{k=l+2}^{n-1}\log A_k\ge \frac{zn}{2})\\
&\ge \frac{c}{n^2}\mathbf{P}(\sum_{k=1}^{l-1}\log A_k\ge \frac{zn}{2})\mathbf{P}(\sum_{k=l+2}^{n-1}\log A_k\le -\frac{zn}{2})
\end{aligned}$$

Applying Cramér's theorem to sums of i.i.d. random variables $\log A_k$, we have

$$\tilde{\mathbb{P}}_0(\tilde\tau_n\wedge\tilde\tau_{-1} > m|A_0\in[a,\frac{1}{a}]) \gtrsim \exp(-n\left(z_1 I(\frac{z}{2z_1}) + (1-z_1)I(\frac{-z}{2(1-z_1)})\right))$$

where $I(x)=\sup_{t\in\mathbb{R}}\{tx-\log\mathbf{E}(A^t)\}$ is the associated rate function. □

Proof of Lemma 3. Replace $I(\frac{-z}{2(1-z_1)})$ using

$$\begin{aligned}
I(-x) &= \sup_{t\in\mathbb{R}}\{-tx-\log\mathbf{E}(A^t)\} = \sup_{t\in\mathbb{R}}\{-tx-\log\mathbf{E}(A^{1-t})\}\\
&= \sup_{s\in\mathbb{R}}\{-(1-s)x-\log\mathbf{E}(A^s)\} = I(x)-x.
\end{aligned}$$

For fixed z, by convexity of the rate function I, the supremum of $-z_1 I(\frac{z}{2z_1}) - (1-z_1)I(\frac{z}{2(1-z_1)})$ is obtained when $z_1=\frac{1}{2}$, we are left to compute

$$\sup_{0<z}\{\frac{\log q_1 - I(z)}{z}+\frac{1}{2}\},$$

clearly, $\frac{\log q_1 - I(z)}{z}\le -t^*$, when z is such that $(t\mapsto \log\mathbf{E}(A^t))'(t^*) = z > 0$, the maximum is obtained. □

Proof of Lemma 4 . Observe that

$$\begin{aligned}
&\widetilde{P}^\omega_{Y_1}(\tilde\tau_y<\tilde\tau_{\overleftarrow{Y_1}})\widetilde{G}^{\tilde\tau_{Y_1}\wedge\tilde\tau_{Y_3}}(y,y) = \widetilde{P}^\omega_{Y_1}(\tilde\tau_y<\tilde\tau_{\overleftarrow{Y_1}}\wedge\tilde\tau_{Y_3})\widetilde{E}^\omega_y\left[\sum_{k=0}^{\tilde\tau_{Y_1}\wedge\tilde\tau_{Y_3}}1_{\{\tilde\eta_k=y\}}\right]\\
\le&\widetilde{P}^\omega_{Y_1}(\tilde\tau_y<\tilde\tau_{\overleftarrow{Y_1}}\wedge\tilde\tau_{Y_3})\widetilde{E}^\omega_y\left[\sum_{k=0}^{\tilde\tau_{\overleftarrow{Y_1}}\wedge\tilde\tau_{Y_3}}1_{\{\tilde\eta_k=y\}}\right] = \widetilde{E}^\omega_{Y_1}\left[\sum_{k=0}^{\tilde\tau_{\overleftarrow{Y_1}}\wedge\tilde\tau_{Y_3}}1_{\{\tilde\eta_k=y\}}\right].
\end{aligned}$$

Obviously,

$$\widetilde{E}^{\omega}_{Y_1}\Big[\sum_{k=0}^{\tilde{\tau}_{\overleftarrow{Y_1}}\wedge\tilde{\tau}_{Y_3}} 1_{\{\widetilde{\eta}_k=y\}}\Big] \le \widetilde{E}^{\omega}_{Y_1}[\tilde{\tau}_{\overleftarrow{Y_1}}\wedge\tilde{\tau}_{Y_3}].$$

This gives us (11).

Moreover, to get (12), we only need to show that for any $0 \le p < m$, we have

$$\widetilde{E}^{\omega}_{p}[\tilde{\tau}_{p-1}\wedge\tilde{\tau}_m] \le 1 + A_pA_{p+1} + A_pA_{p+1}\widetilde{E}^{\omega}_{p+1}[\tilde{\tau}_p\wedge\tilde{\tau}_m]. \tag{39}$$

In fact, since $0 \le \lambda \le 1$, (39) implies that

$$\widetilde{E}^{\omega}_{p}[\tilde{\tau}_{p-1}\wedge\tilde{\tau}_m]^{\lambda} \le 1 + (A_pA_{p+1})^{\lambda} + (A_pA_{p+1})^{\lambda}\widetilde{E}^{\omega}_{p+1}[\tilde{\tau}_p\wedge\tilde{\tau}_m]^{\lambda}.$$

applying this inequality a few times along the interval $[\![Y_1, Y_3]\!]$, we obtain (12). It remains to show (39). Observe that

$$\begin{aligned}
&\widetilde{E}^{\omega}_{p}[\tilde{\tau}_{p-1}\wedge\tilde{\tau}_m] = \widetilde{\omega}(p,p-1) + \widetilde{\omega}(p,p+1)(1+\widetilde{E}^{\omega}_{p+1}[\tilde{\tau}_{p-1}\wedge\tilde{\tau}_m])\\
&= 1 + \widetilde{\omega}(p,p+1)\widetilde{E}^{\omega}_{p+1}[\tilde{\tau}_{p-1}\wedge\tilde{\tau}_m]\\
&= 1 + \widetilde{\omega}(p,p+1)\Big(\widetilde{E}^{\omega}_{p+1}[\tilde{\tau}_m;\tilde{\tau}_m<\tilde{\tau}_p] + \widetilde{E}^{\omega}_{p+1}[\tilde{\tau}_p;\tilde{\tau}_p<\tilde{\tau}_m] + \widetilde{P}^{\omega}_{p+1}(\tilde{\tau}_p<\tilde{\tau}_m)\widetilde{E}^{\omega}_{p}[\tilde{\tau}_{p-1}\wedge\tilde{\tau}_m]\Big).
\end{aligned}$$

It follows that

$$\begin{aligned}
\widetilde{E}^{\omega}_{p}[\tilde{\tau}_{p-1}\wedge\tilde{\tau}_m] &= \frac{1+\widetilde{\omega}(p,p+1)\widetilde{E}^{\omega}_{p+1}[\tilde{\tau}_p\wedge\tilde{\tau}_m]}{1-\widetilde{\omega}(p,p+1)\widetilde{P}^{\omega}_{p+1}(\tilde{\tau}_p<\tilde{\tau}_m)}\\
&= \frac{1+\widetilde{\omega}(p,p+1)\widetilde{E}^{\omega}_{p+1}[\tilde{\tau}_p\wedge\tilde{\tau}_m]}{\widetilde{\omega}(p,p-1)+\widetilde{\omega}(p,p+1)\widetilde{P}^{\omega}_{p+1}(\tilde{\tau}_m<\tilde{\tau}_p)} \le \frac{1+\widetilde{\omega}(p,p+1)\widetilde{E}^{\omega}_{p+1}[\tilde{\tau}_p\wedge\tilde{\tau}_m]}{\widetilde{\omega}(p,p-1)}.
\end{aligned}$$

Therefore,

$$\widetilde{E}^{\omega}_{p}[\tilde{\tau}_{p-1}\wedge\tilde{\tau}_m] \le (1+A_pA_{p+1}) + A_pA_{p+1}\widetilde{E}^{\omega}_{p+1}[\tilde{\tau}_p\wedge\tilde{\tau}_m].$$

□

Proof of Lemma 5. Recall that $\mathbf{E}[A^t] < \infty$ for any $t \in \mathbb{R}$. By Hölder's inequality, it suffices to show that there exists some $\delta' > 0$ such that for all n large enough,

$$\mathbf{E}\Big[\Big(\tilde{E}^{\omega}_{0}[\tilde{\tau}_{-1}\wedge\tilde{\tau}_n]\Big)^{\lambda(1+\delta')}\Big] \le (q_1+\delta)^{-n}. \tag{40}$$

It remains to prove (40). In fact, we only need to show that for $1 > \lambda' = \lambda(1+\delta) > 0$,

$$\limsup_{n\to\infty} \frac{\log \mathbf{E}\Big[\Big(\tilde{E}^{\omega}_{0}[\tilde{\tau}_{-1}\wedge\tilde{\tau}_n]\Big)^{\lambda'}\Big]}{n} \le \psi(\lambda'+1/2) \tag{41}$$

where $\psi(t) = \log \mathbf{E}(A^t)$. One therefore sees that if $t^* - 1/2 > \lambda'$, then $\psi(\lambda'+1/2) < \psi(t^*) = -\log q_1$. To show (41), recall that for any $0 \le i \le n-1$,

$$\begin{aligned}
\widetilde{G}^{\tilde{\tau}_{-1}\wedge\tilde{\tau}_n}(i,i) &= \widetilde{E}^{\omega}_{i}\Big[\sum_{k=0}^{\tilde{\tau}_{-1}\wedge\tilde{\tau}_n} 1_{\eta=i}\Big]\\
&= \frac{1}{1-\widetilde{\omega}(i,i-1)\widetilde{P}_{i-1}(\tilde{\tau}_i<\tilde{\tau}_{-1})-\widetilde{\omega}(i,i+1)\widetilde{P}_{i+1}(\tilde{\tau}_i<\tilde{\tau}_n)}.
\end{aligned}$$

Then, $\widetilde{E}_0^\omega[\tilde{\tau}_{-1}\wedge\tilde{\tau}_n] = 1 + \sum_{i=0}^{n-1}\widetilde{P}_0^\omega(\tilde{\tau}_i<\tilde{\tau}_{-1})\widetilde{G}^{\tilde{\tau}_{-1}\wedge\tilde{\tau}_n}(i,i)$ implies that

$$\widetilde{E}_0^\omega[\tilde{\tau}_{-1}\wedge\tilde{\tau}_n] = 1 + \sum_{i=0}^{n-1}\frac{\widetilde{P}_0^\omega(\tilde{\tau}_i<\tilde{\tau}_{-1})}{\widetilde{\omega}(i,i-1)\widetilde{P}_{i-1}^\omega(\tilde{\tau}_{-1}<\tilde{\tau}_i)+\widetilde{\omega}(i,i+1)\widetilde{P}_{i+1}^\omega(\tilde{\tau}_n<\tilde{\tau}_i)}.$$

Recall that by (38), if $S_i := \sum_{j=1}^i -\log(A_{j-1}A_j)$ for $i\geq 1$ and $S_0=0$, then

$$\begin{aligned}\widetilde{P}_0^\omega(\tilde{\tau}_i<\tilde{\tau}_{-1}) &= \frac{1}{\sum_{k=0}^i e^{S_k}}\\ \widetilde{P}_{i-1}^\omega(\tilde{\tau}_{-1}<\tilde{\tau}_i) &= \frac{e^{S_i}}{\sum_{k=0}^i e^{S_k}}\\ \widetilde{P}_{i+1}^\omega(\tilde{\tau}_n<\tilde{\tau}_i) &= \frac{1}{\sum_{k=i+1}^n e^{S_k-S_{i+1}}}\end{aligned}$$

It is immediate that

$$\begin{aligned}\frac{\widetilde{P}_0^\omega(\tilde{\tau}_i<\tilde{\tau}_{-1})}{\widetilde{\omega}(i,i-1)\widetilde{P}_{i-1}^\omega(\tilde{\tau}_{-1}<\tilde{\tau}_i)+\widetilde{\omega}(i,i+1)\widetilde{P}_{i+1}^\omega(\tilde{\tau}_n<\tilde{\tau}_i)} =& \frac{\frac{1}{\sum_{k=0}^i e^{S_k}}}{\frac{1}{1+A_iA_{i+1}}\frac{e^{S_i}}{\sum_{k=0}^i e^{S_k}}+\frac{A_iA_{i+1}}{1+A_iA_{i+1}}\frac{1}{\sum_{k=i+1}^n e^{S_k-S_{i+1}}}}\\ \leq& \frac{1}{\frac{1}{1+A_iA_{i+1}}\frac{e^{S_i}}{\sum_{k=0}^i e^{S_k}}+\frac{A_iA_{i+1}}{1+A_iA_{i+1}}\frac{1}{\sum_{k=i+1}^n e^{S_k-S_{i+1}}}}.\end{aligned}$$

Let $X_k=-\log A_k$. For any $0\leq i\leq n$, define

$$\begin{aligned}H_i(-X) &:= \max_{0\leq j\leq i}(-X_j-X_{j+1}-\cdots-X_{i-1})\\ H_{n-i-1}(X) &:= \max_{i+2\leq j\leq n}(X_{i+2}+\cdots+X_j)\end{aligned}$$

Note that

$$S_k-S_i\leq 2H_i(-X)+(-X_i)_+, \forall 0\leq k\leq i,$$

and that

$$S_k-S_{i+1}\leq 2H_{n-i-1}(X)+(X_{i+1})_+, \forall i+1\leq k\leq n.$$

Then,

$$\frac{1}{1+A_iA_{i+1}}\frac{e^{S_i}}{\sum_{k=0}^i e^{S_k}}\geq\frac{1}{1+A_iA_{i+1}}\frac{1}{(1+i)e^{2H_i(-X)+(-X_i)_+}}\geq\frac{1}{n(A_i+1)(1+A_iA_{i+1})}e^{-2H_i(-X)}.$$

Similarly,

$$\frac{A_iA_{i+1}}{1+A_iA_{i+1}}\frac{1}{\sum_{k=i+1}^n e^{S_k-S_{i+1}}}\geq\frac{(A_{i+1}\wedge 1)A_iA_{i+1}}{n(1+A_iA_{i+1})}e^{-2H_{n-i-1}(X)}.$$

So,

$$\begin{aligned}&\frac{1}{1+A_iA_{i+1}}\frac{e^{S_i}}{\sum_{k=0}^i e^{S_k}}+\frac{A_iA_{i+1}}{1+A_iA_{i+1}}\frac{1}{\sum_{k=i+1}^n e^{S_k-S_{i+1}}}\\ &\qquad\geq\frac{1}{n(A_i+1)(1+A_iA_{i+1})}e^{-2H_i(-X)}+\frac{(A_{i+1}\wedge 1)A_iA_{i+1}}{n(1+A_iA_{i+1})}e^{-2H_{n-i-1}(X)}\\ &\qquad\quad\geq\frac{1}{n}\Big(\frac{1}{(A_i\vee 1)(1+A_iA_{i+1})}\wedge\frac{(A_{i+1}\wedge 1)A_iA_{i+1}}{1+A_iA_{i+1}}\Big)e^{-2H_i(-X)}\vee e^{-2H_{n-i-1}(X)}.\end{aligned}$$

This implies that

$$\begin{aligned}&\frac{\widetilde{P}_0^\omega(\tilde\tau_i<\tilde\tau_{-1})}{\widetilde\omega(i,i-1)\widetilde{P}_{i-1}^\omega(\tilde\tau_{-1}<\tilde\tau_i)+\widetilde\omega(i,i+1)\widetilde{P}_{i+1}^\omega(\tilde\tau_n<\tilde\tau_i)}\\&\le n\Big((A_i\vee 1)(1+A_iA_{i+1})+\frac{1+A_iA_{i+1}}{(A_{i+1}\wedge 1)A_iA_{i+1}}\Big)e^{2H_i(-X)\wedge H_{n-i-1}(X)}.\end{aligned}$$

Thus, for any $\lambda\le 1$, $n\ge 2$,

$$\widetilde{E}_0^\omega[\tilde\tau_{-1}\wedge\tilde\tau_n]^\lambda\lesssim n+n^2\sum_{i=0}^{n-1}\Big((A_i\vee 1)(1+A_iA_{i+1})+\frac{1+A_iA_{i+1}}{(A_{i+1}\wedge 1)A_iA_{i+1}}\Big)^\lambda e^{2\lambda H_i(-X)\wedge H_{n-i-1}(X)}$$

By independence,

$$\mathbf{E}\widetilde{E}_0^\omega[\tilde\tau_{-1}\wedge\tilde\tau_n]^\lambda\lesssim n+n^3\max_{0\le i\le n-1}\mathbf{E}[e^{2\lambda H_i(-X)\wedge H_{n-i-1}(X)}]\tag{42}$$

Recall that $\psi(\lambda)=\log\mathbf{E}[A^\lambda]$ and $\mathscr{S}_k=-\sum_{i=1}^k\log A_i$. Let $t>0$, for $i\ge 1$, $x>0$,

$$\begin{aligned}\mathbf{P}(H_i(-X)\ge xi)&\le\mathbf{P}(\max_{0\le k\le i}[-t\mathscr{S}_k-\psi(t)k]\ge xti-\psi(t)i)\\&\le\mathbf{P}(\max_{0\le k\le i}e^{-t\mathscr{S}_k-\psi(t)k}\ge e^{(xt-\psi(t))i})\\&\le e^{-(xt-\psi(t))i},\end{aligned}\tag{43}$$

where the last inequality stem from Doob's maximal inequality and the fact that $(e^{-t\mathscr{S}_j-\psi(t)j})_j$ is a martingale. Since $x\ge\mathbf{E}(\log A)$, $I(x)=\sup_{t>0}\{tx-\psi(t)\}$, we have

$$\mathbf{P}(H_i(-X)\ge xi)\le e^{-I(x)i}.\tag{44}$$

Similarly, for any $j\ge 1$ and $x>\mathbf{E}[-\log A]$.

$$\begin{aligned}\mathbf{P}(H_j(X)\ge xj)&\le\mathbf{P}(\max_{0\le k\le j}[t\mathscr{S}_k-\psi(-t)k]\ge xtj-\psi(-t)j)\\&\le\mathbf{P}(\max_{0\le k\le j}e^{t\mathscr{S}_k-\psi(-t)k}\ge e^{(xt-\psi(-t))j})\\&\le e^{-(xt-\psi(-t))j},\end{aligned}\tag{45}$$

which implies that

$$\mathbf{P}(H_j(X)\ge xj)\le e^{-I(-x)j}.\tag{46}$$

Further, for $0<x<\mathbf{E}[-\log A]$, one sees that by Cramér's theorem,

$$\begin{aligned}\mathbf{P}(H_j(X)\le xj)&\le\mathbf{P}(X_1+\cdots+X_j\le xj)\\&=\mathbf{P}(-X_1-\cdots-X_j\ge -xj)\le e^{-I(-x)j}.\end{aligned}\tag{47}$$

Take $\eta>0$. In (42), we can replace $H_i(-X)\wedge H_{n-i-1}(X)$ by $H_i(-X)\wedge H_{n-i-1}(X)\wedge K\eta n$ with some $K\ge 1$ large enough. In fact,

$$\begin{aligned}\mathbf{E}[e^{2\lambda H_i(-X)\wedge H_{n-i-1}(X)}]\le&\underbrace{\mathbf{E}[e^{2\lambda H_i(-X)\wedge H_{n-i-1}(X)};H_i(-X)\vee H_{n-i-1}(X)\le K\eta n]}_{\Xi_K^-(i)}\\&+\underbrace{\mathbf{E}[e^{2\lambda H_i(-X)\wedge H_{n-i-1}(X)};H_i(-X)\vee H_{n-i-1}(X)\ge K\eta n]}_{=:\Xi_K^+(i)}.\end{aligned}$$

Observe that

$$\begin{aligned}\Xi_K^+(i) \leq &\mathbf{E}[e^{2\lambda H_i(-X)}; H_i(-X) \geq K\eta n] + \mathbf{E}[e^{2\lambda H_{n-i-1}(X)}; H_{n-i-1}(X) \geq K\eta n] \\ =: &\Xi_1 + \Xi_2\end{aligned}$$

Let us bound Ξ_1,

$$\begin{aligned}\Xi_1 =&\mathbf{E}\int_{-\infty}^{H_i(-X)} 2\lambda e^{2\lambda x}\mathbf{1}_{H_i(-X)\geq K\eta n}dx = \int_{\mathbb{R}} 2\lambda e^{2\lambda x}\mathbf{P}(H_i(-X) \geq K\eta n \vee x)dx \\ =&\int_{-\infty}^{K\eta n} 2\lambda e^{2\lambda x}dx\mathbf{P}(H_i(-X) \geq K\eta n) + \int_{K\eta n}^{\infty} 2\lambda e^{2\lambda x}\mathbf{P}(H_i(-X) \geq x)dx \\ =&e^{2\lambda K\eta n}\mathbf{P}(H_i(-X) \geq K\eta n) + \int_K^{\infty} 2\lambda\eta n e^{2\lambda t\eta n}\mathbf{P}(H_i(-X) \geq t\eta n)dt\end{aligned}$$

By applying (43), one sees that for any $0 \leq i \leq n-1$ and $\mu = 3 > 2\lambda$,

$$\begin{aligned}\Xi_1 \leq &e^{2\lambda K\eta n}e^{-\mu K\eta n+\psi(\mu)i} + \int_K^{\infty} 2\lambda\eta n e^{2\lambda t\eta n}e^{-\mu t\eta n+\psi(\mu)i}dt \\ \leq &e^{-K\eta n+\psi(3)n} + 2\lambda e^{\psi(3)n}\int_K^{\infty} \eta n e^{-t\eta n}dt \\ \leq &3e^{-K\eta n+\psi(3)n},\end{aligned}$$

which is less than 1 when we choose K large enough. Similarly, we can show that for any $i \leq n-1$,

$$\Xi_2 \leq 1,$$

for K large enough. Consequently, (42) becomes that

$$\mathbf{E}\widetilde{E}_0^{\omega}[\tilde{\tau}_{-1} \wedge \tilde{\tau}_n]^{\lambda} \lesssim 3n^3 + n^3 \max_{0\leq i\leq n-1} \Xi_K^-(i). \tag{48}$$

It remains to bound $\Xi_K^-(i)$. Take sufficiently small $\varepsilon > 0$ and let $L = \lfloor \frac{1}{\varepsilon} \rfloor$. For any i such that $l_1\lfloor \varepsilon n\rfloor \leq i < (l_1+1)\lfloor \varepsilon n\rfloor$ and $l_2\lfloor \varepsilon n\rfloor \leq n-i-1 < (l_2+1)\lfloor \varepsilon n\rfloor$ with $0 \leq l_1, l_2 \leq L$, we have

$$\begin{aligned}\Xi_K^-(i) \leq &\sum_{0\leq k_1,k_2\leq K} e^{2\lambda k_1\wedge k_2\eta n+2\lambda\eta n}\mathbf{P}(k_1\eta n \leq H_i(-X) < (k_1+1)\eta n)\mathbf{P}(k_2\eta n \leq H_{n-i-1}(X) < (k_2+1)\eta n) \\ \leq &\sum_{0\leq k_1,k_2\leq K} e^{2\lambda k_1\wedge k_2\eta n+2\lambda\eta n}\mathbf{P}(H_i(-X) \geq k_1\eta n)\mathbf{P}(k_2\eta n \leq H_{n-i-1}(X) < (k_2+1)\eta n).\end{aligned}$$

By (44), we have

$$\mathbf{P}(H_i(-X) \geq k_1\eta n) \leq e^{-I(x_1)i}$$

where x_1 is the point in $[\frac{k_1\eta n}{(l_1+1)\lfloor\varepsilon n\rfloor}, \frac{k_1\eta n}{l_1\lfloor\varepsilon n\rfloor}]$ where I reaches the minimum in this interval. By large deviation estimates (46) (47), we have

$$\mathbf{P}(k_2\eta n \leq H_{n-i-1}(X) < (k_2+1)\eta n) \leq e^{-I(x_2)(n-i)}$$

where x_2 is the point in $[\frac{k_1\eta n}{(l_2+1)\lfloor\varepsilon n\rfloor}, \frac{(k_2+1)\eta n}{l_2\lfloor\varepsilon n\rfloor}]$ where I reaches the minimum in this interval. Therefore,

$$\Xi_K^-(i) \leq \sum_{0\leq k_1,k_2\leq K} e^{2\lambda k_1\wedge k_2\eta n+2\lambda\eta n}e^{-I(x_1)l_1\lfloor\varepsilon n\rfloor}e^{-I(-x_2)l_2\lfloor\varepsilon n\rfloor}$$

Taking maximum over all l_1, l_2, k_1, k_2 yields that

$$\mathbf{E}\widetilde{E}_0^{\omega}[\tilde{\tau}_{-1}\wedge\tilde{\tau}_n]^{\lambda} \lesssim 3n^2+n^2K^2 \max_{l_1,l_2,k_1,k_2} \exp\{2\lambda k_1\wedge k_2\eta n+2\lambda\eta n-I(x_1)l_1\lfloor\varepsilon n\rfloor-I(-x_2)l_2\lfloor\varepsilon n\rfloor\}. \quad (49)$$

Observe that

$$\begin{aligned}&2\lambda k_1 \wedge k_2\eta n + 2\lambda\eta n - I(x_1)l_1\lfloor\varepsilon n\rfloor - I(-x_2)l_2\lfloor\varepsilon n\rfloor\\ \le&2\lambda(x_1l_1 \wedge x_2l_2)\lfloor\varepsilon n\rfloor - I(x_1)l_1\lfloor\varepsilon n\rfloor - I(-x_2)l_2\lfloor\varepsilon n\rfloor + 3\lambda\eta n.\end{aligned}$$

Define

$$L(\lambda) := \sup_{\mathcal{D}}\{\Big(x_1z_1 \wedge x_2z_2\Big)\lambda - I(x_1)z_1 - I(-x_2)z_2\},$$

where $\mathcal{D} := \{x_1, x_2, z_1, z_2 \ge 0, z_1 + z_2 \le 1\}$.

By Lemma 8.1 in [1], one concludes that

$$\limsup_{n\to\infty} \frac{\log \mathbf{E}\widetilde{E}_0^{\omega}[\tilde{\tau}_{-1} \wedge \tilde{\tau}_n]^{\lambda}}{n} \le L(2\lambda) = \psi(\frac{1 + 2\lambda}{2}).$$

□

B Some observations on random walks on random trees

Proof of Lemma 10. As $\beta(x)$ is identically distributed under $\mathbb{P}$,

$$\begin{aligned}\mathbb{E}_{\rho}(\sum_{|x|=n} \mathbb{1}_{\tau_x<\infty})\mathbb{E}(\beta) &= \mathbb{E}[\sum_{|x|=n} P_{\rho}^{\omega,T}(\tau_x < \infty)]\mathbb{E}(\beta)\\ &= \mathbb{E}\left(\sum_{|x|=n} \mathbb{E}^T(P_{\rho}^{\omega,T}(\tau_x < \infty))\mathbb{E}^T(\beta(x))\right).\end{aligned}$$

$P_{\rho}^{\omega,T}(\tau_x < \infty)$ is an increasing function of A_x since

$$\begin{aligned}P_{\rho}^{\omega,T}(\tau_x < \infty) &= P_{\rho}^{\omega,T}(\tau_{\overleftarrow{x}} < \infty)\left(\sum_{k\ge0} P_{\overleftarrow{x}}^{\omega,T}(\tau_{\overleftarrow{x}}^* < \min(\tau_x, \infty))^k\right) p(\overleftarrow{x}, x)\\ &= \frac{P_{\rho}^{\omega,T}(\tau_{\overleftarrow{x}} < \infty)}{1 - P_{\overleftarrow{x}}^{\omega,T}(\tau_{\overleftarrow{x}}^* < \min(\tau_x, \infty))} \frac{A_{\overleftarrow{x}}A_x}{1 + A_{\overleftarrow{x}}B_{\overleftarrow{x}}},\end{aligned}$$

recall that $\beta(x)$ is also an increasing function of A_x, moreover, conditionally on A_x, $P_{\rho}^{\omega,T}(\tau_x < \infty)$ and $\beta(x)$ are independent, thus by FKG inequality,

$$\begin{aligned}\mathbb{E}^T(P_{\rho}^{\omega,T}(\tau_x < \infty)\beta(x)) &= \mathbb{E}^T(\mathbb{E}^T(P_{\rho}^{\omega,T}(\tau_x < \infty)\beta(x)|A_x))\\ &= \mathbb{E}^T(\mathbb{E}^T(P_{\rho}^{\omega,T}(\tau_x < \infty)|A_x)\mathbb{E}^T(\beta(x)|A_x))\\ &\ge \mathbb{E}^T(P_{\rho}^{\omega,T}(\tau_x < \infty))\mathbb{E}^T(\beta(x))\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{E}\left(\sum_{|x|=n} \mathbb{E}^T(P_{\rho}^{\omega,T}(\tau_x < \infty))\mathbb{E}^T(\beta(x))\right) &\le \mathbb{E}\left(\sum_{|x|=n} \mathbb{E}^T(P_{\rho}^{\omega,T}(\tau_x < \infty)\beta(x))\right)\\ &= \mathbb{E}\left(\sum_{|x|=n} P_{\rho}^{\omega,T}(\tau_x < \infty)\beta(x)\right)\end{aligned}$$

For any GW tree and any trajectory on the tree, there is at most one regeneration time at the n-th generation, therefore,

$$\sum_{|x|=n} \mathbb{1}_{\tau_x<\infty,\ \eta_k\neq \overleftarrow{x}, \forall k>\tau_x} \leq 1$$

By taking expectation w.r.t. $E_\rho^{\omega,T}$ and using the Markov property at τ_x,

$$\sum_{|x|=n} P_\rho^{\omega,T}(\tau_x<\infty)\beta(x) \leq 1$$

Whence

$$\mathbb{E}(\sum_{|x|=n} \mathbb{1}_{\tau_x<\infty})\mathbb{E}(\beta) \leq 1$$

By transient assumption it suffices to take $c_{11} = \frac{1}{\mathbb{E}(\beta)} < \infty$. □

Proof of Lemma 11 and Corollary 3. Let $T_i,\ i \geq 1$ be independent copies of GW tree with offspring distribution (q), each endowed with independent environment $(\omega_x, x \in T_i)$. Let $\rho^{(i)}$ be the root of T_i. In such setting, $\beta(\rho^{(i)}),\ i \geq 1$ are i.i.d. sequence with common distribution β.

For each T_i, take the left most infinite ray, denoted $v_0^{(i)} = \rho^{(i)}, v_1^{(i)}, \ldots, v_n^{(i)}, \ldots$ Let $\Omega(x) = \{y \neq x;\ \overleftarrow{x} = \overleftarrow{y}\}$ be the set of all brothers of x. Fix some constant C, define

$$R_i = \inf\{n \geq 1;\ \exists z \in \Omega(v_n^{(i)}),\ \frac{1}{A_z\beta(z)} \leq C\}.$$

By Equation (15),

$$\frac{1}{\beta(v_{R_i-1}^{(i)})} \leq 1 + \frac{1}{A_{v_{R_i-1}^{(i)}} A_z \beta(z)} \leq 1 + \frac{C}{A_{v_{R_i-1}^{(i)}}}.$$

Also R_i and $\{A_{v_n^{(i)}},\ n \geq 0\}$ are independent under Q. By iteration,

$$\begin{aligned}
\frac{1}{\beta(\rho^{(i)})} &\leq 1 + \frac{1}{A_{v_0^{(i)}} A_{v_1^{(i)}} \beta(v_1^{(i)})} \leq 1 + \frac{1}{A_{v_0^{(i)}} A_{v_1^{(i)}}}\left(1 + \frac{1}{A_{v_1^{(i)}} A_{v_2^{(i)}} \beta(v_2^{(i)})}\right) \\
&\leq \cdots \\
&\leq 1 + \sum_{k=1}^{R_i-1} \frac{1}{A_{v_0^{(i)}} A_{v_k^{(i)}}} \prod_{j=1}^{k-1} A_{v_j^{(i)}}^{-2} + \frac{C}{A_{v_0^{(i)}}} \prod_{l=1}^{R_i-1} A_{v_l^{(i)}}^{-2}.
\end{aligned}$$

For any $n \geq 0$, denote

$$\mathcal{C}(n) = 1 + \sum_{k=1}^{n} \frac{1}{A_{v_0^{(i)}} A_{v_k^{(i)}}} \prod_{j=1}^{k-1} A_{v_j^{(i)}}^{-2} + \frac{C}{A_{v_0^{(i)}}} \prod_{l=1}^{n} A_{v_l^{(i)}}^{-2}. \tag{50}$$

Thus $\dfrac{1}{\beta(\rho^{(i)})} \leq \mathcal{C}(R_i - 1)$, note also that, since $\xi_2 = \mathbf{E}(A^{-2}) = 1 + \frac{3}{c^2} + \frac{3}{c^4}$, $E(\mathcal{C}(n)) \leq c_{34}\xi_2^{n+1}$. Therefore, for any $K \geq 1$,

$$\frac{1}{\sum_{i=1}^K \beta(\rho^{(i)})} \leq \mathcal{C}(\min_{1\leq i\leq K} R_i - 1).$$

Taking expectation under $\mathbb{P}$ yields (as R_i i.i.d. let R be a r.v. with the common distribution)

$$\begin{aligned}\mathbb{E}(\frac{1}{\sum_{i=1}^K \beta(\rho^{(i)})}) &\leq \mathbb{E}(\mathbb{E}(\mathcal{C}(\min_{1\leq i\leq K} R_i - 1)|R_i;\ 1\leq i\leq K)) \\ &\leq c_{34}\mathbb{E}(\xi_2^{\min_{1\leq i\leq K} R_i}) \leq c_{34}\sum_{n=0}^{\infty} \xi_2^{n+1}\mathbb{P}(R\geq n+1)^K \\ &\leq c_{34}\sum_{n\geq 0}\xi_2^{n+1}\mathbb{E}(\delta_C^{\sum_{k=0}^{n-1}(d(v_k)-2)})^K\end{aligned}$$

where $\delta_C = \mathbb{P}(\frac{1}{A_\rho\beta_\rho} > C)$. Let $f(s) = \sum_{k\geq 1} q_k s^k$, as $f(s)/s \downarrow q_1$ as $s\downarrow 0$, for any $\varepsilon > 0$, we can take C large enough to ensure $\frac{f(\delta_C)}{\delta_C} \leq q_1(1+\varepsilon)$, thus

$$\mathbb{E}(\frac{1}{\sum_{i=1}^K \beta(\rho^{(i)})}) \leq c_{34}\sum_{n\geq 0}\xi_2^{n+1}(\frac{f(\delta_C)}{\delta_C})^{nK} \leq c_{34}\sum_{n\geq 0}\xi_2^{n+1}(q_1(1+\varepsilon))^{nK}.$$

Now take ε such that $q_1(1+\varepsilon) < 1$, then take K large enough such that $\xi_2(q_1(1+\varepsilon))^K < 1$ leads to

$$\mathbb{E}(\frac{1}{\sum_{i=1}^K \beta(\rho^{(i)})}) < c_{12} < \infty$$

Similarly, the following also holds

$$\mathbb{E}(\frac{1}{\sum_{i=1}^K A_{\rho^{(i)}}\beta(\rho^{(i)})}) < c_{12} < \infty.$$

In particular, if $q_1\zeta_2 < 1$, we can take $K = 1$ and obtained Further, it follows from (50) and Chauchy-Schwartz inequality that

$$\mathcal{C}(n)^2 \leq (n+2)\Big(1 + \sum_{k=1}^n \frac{1}{A^2_{v_0^{(i)}}A^2_{v_k^{(i)}}}\prod_{j=1}^{k-1} A^{-4}_{v_j^{(i)}} + \frac{C}{A_{v_0^{(i)}}}\prod_{l=1}^n A^{-4}_{v_l^{(i)}}\Big).$$

Thus,

$$\mathbb{E}[\mathcal{C}^2(n)] \leq c_{35}(n+2)\xi_4^{n+1}.$$

As soon as $\zeta_4 < \infty$, the previous argument works again to conclude that for K large enough,

$$\mathbb{E}(\frac{1}{\sum_{i=1}^K \beta^2(\rho^{(i)})}) + \mathbb{E}(\frac{1}{\sum_{i=1}^{K_0} A^2_{\rho^{(i)}}\beta^2(\rho^{(i)})}) < c_{13} < \infty.$$

□

References

[1] Elie Aidékon. Transient random walks in random environment on a galton–watson tree. *Probability Theory and Related Fields*, 142(3-4):525–559, 2008.

[2] Omer Angel, Nicholas Crawford, and Gady Kozma. Localization for linearly edge reinforced random walks. *Duke Mathematical Journal*, 163(5):889–921, 2014.

[3] Anne-Laure Basdevant, Arvind Singh, et al. Continuous-time vertex reinforced jump processes on galton–watson trees. *The Annals of Applied Probability*, 22(4):1728–1743, 2012.

[4] John D Biggins. Martingale convergence in the branching random walk. *Journal of Applied Probability*, pages 25–37, 1977.

[5] Andrea Collevecchio. Limit theorems for vertex-reinforced jump processes on regular trees. *Electron. J. Probab*, 14(66):1936–1962, 2009.

[6] Burgess Davis and Stanislav Volkov. Continuous time vertex-reinforced jump processes. *Probability theory and related fields*, 123(2):281–300, 2002.

[7] Burgess Davis and Stanislav Volkov. Vertex-reinforced jump processes on trees and finite graphs. *Probability theory and related fields*, 128(1):42–62, 2004.

[8] Margherita Disertori and Tom Spencer. Anderson localization for a supersymmetric sigma model. *Communications in Mathematical Physics*, 300(3):659–671, 2010.

[9] Margherita Disertori, Tom Spencer, and Martin R Zirnbauer. Quasi-diffusion in a 3d supersymmetric hyperbolic sigma model. *Communications in Mathematical Physics*, 300(2):435–486, 2010.

[10] Russell Lyons. A simple path to biggins' martingale convergence for branching random walk. In *Classical and modern branching processes*, pages 217–221. Springer, 1997.

[11] Russell Lyons and Robin Pemantle. Random walk in a random environment and first-passage percolation on trees. *The Annals of Probability*, pages 125–136, 1992.

[12] Russell Lyons and Yuval Peres. Probability on trees and networks. book in progress, 2005.

[13] Christophe Sabot and Pierre Tarres. Edge-reinforced random walk, vertex-reinforced jump process and the supersymmetric hyperbolic sigma model. *arXiv preprint arXiv:1111.3991*, 2011.

[14] Christophe Sabot and Pierre Tarres. Ray-knight theorem: a short proof. *arXiv preprint arXiv:1311.6622*, 2013.